

UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE PATENT TRIAL AND APPEAL BOARD

Google LLC
“Google”
Petitioner

v.

Advanced Coding Technologies, LLC
“ACT”
Patent Owner

INTER PARTES REVIEW OF U.S. PATENT NO. 9,042,448

DECLARATION OF DR. CLIFFORD READER

I declare that all statements made in this declaration on my own knowledge are true and that all statements made on information and belief are believed to be true, and further, that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

By: c reader

Clifford Reader, Ph.D.

Contents

I.	Professional Background.....	9
A.	Summary	9
B.	Education.....	9
C.	Work Experience.....	10
D.	Standardization.....	11
E.	Intellectual Property Rights.....	13
F.	Curriculum Vitae.....	14
II.	Digital Video Technologies.....	15
A.	Analog Video	15
B.	Digital Video	16
C.	Image Restoration; Superresolution.....	20
D.	Image Warping.....	27
E.	Video Coding	30
F.	Loop Filters	35
G.	Video Coding Standards.....	38
1.	Background.....	38
2.	The H.261 Standard	40
3.	The MPEG1 Standard.....	43
4.	The MPEG2 Standard.....	46
5.	MPEG Scalability	47
6.	The H.263 Standard	52

7.	The H.264 Standard	61
III.	Relevant Legal Standards	70
IV.	Person of Ordinary Skill in the Art.....	72
V.	Overview of the Subject Patent 9,042,448	73
A.	Overview	80
1.	Resolution; Claims 1-3	80
2.	Combination of Processes; Claims 1-3	82
3.	Enlargement; Claims 1-3	83
4.	Resolution Conversion; Claims 1-3	83
5.	Superresolution enlargement ... to create a set of super- resolution enlarged [decoded] pictures with a resolution higher than the standard resolution; Claims 1-3	84
6.	Sequence, sub-sequence (of pictures), Set (of pictures); Claims 1-3	85
7.	Encoding; Claims 1-3	87
B.	The patent specification	88
1.	Field of Art.....	88
2.	Background - Problem statement, Admitted Prior Art	88
3.	Summary of the disclosed subject matter	89
4.	Comments	102
C.	Prosecution History of the '448 Patent	105
VI.	Summary of the Applied Prior Art	106
A.	Overview of Phek.....	106
B.	Overview of Martins	109

C.	Overview of He	111
VII.	Claim Construction.....	115
VIII.	Ground 1 – Claims 1-3 are Obvious under 35 U.S.C. § 103 based on Phek in view of Martins, and He	116
A.	Overview of the Phek-Martins-He Combination	116
1.	Obvious to Generate Phek’s Extended Layer Image Sequence Target Images by Super-Resolution Processing.....	116
2.	Obvious to Upscale Phek’s Reference Pictures and Encoding Target Pictures to the Same Spatial Resolution	124
3.	Obvious Based on He to Selectively Apply Phek’s/Martin’s Super-Resolution (and Upscaling) Loop Filter.....	129
B.	Relevance of the Phek- Martins-He Combination to the Claims.....	134
1.	Claim 1	134
2.	Claim 2	166
3.	Claim 3	168
IX.	Ground 2 – Claims 1-3 are Obvious under 35 U.S.C. § 103 based on Sakazume in view of Martins and He.....	170
A.	Sakazume.....	170
B.	Combination of Sakazume, Martins, and He	173
1.	Obvious Based on Martins to Use Super-resolution for Upscaling Base-layer Resolution Video	173
2.	Obvious Based on Martins and He to Selectively Use Super- Resolution or Interpolated Images	178
C.	Relevance of the Sakazume-Martins-He Combination to the Claims.....	181
1.	Claim 1	181

2.	Claim 2	193
3.	Claim 3	195
X.	Additional Remarks	197

I, Clifford Reader, Ph.D., do hereby declare:

1. I am making this declaration at the request of Google in the matter of INTER PARTES REVIEW OF U.S. PATENT NO. 9,042,448, “the ’448 Patent.”

2. I am being compensated for my work in this matter at my standard hourly rate of \$850 for consulting services. My compensation in no way depends on the outcome of this proceeding or the content of my testimony.

3. In preparing this Declaration, I considered the following materials:

- | | |
|---------|---|
| Ex.1001 | U.S. Patent No. 9,042,448 to Sakazume (“the ’448 Patent”) |
| Ex.1002 | Excerpts from the Prosecution History of the ’448 Patent (“the ’448 Patent File History”) |
| Ex.1005 | English Translation of Japanese Patent Publication No. 2007316161 A with Translation Certificate (“Phek”) |
| Ex.1006 | Segall et al., <i>Spatial Scalability Within the H.264/AVC Scalable Video Coding Extension</i> ; IEEE Transactions on Circuits & Systems for Video Tech., Vol. 17, No. 9, September 2007 (“Segall”) |
| Ex.1007 | Martins et al., <i>A Unified Approach to Restoration, Deinterlacing and Resolution Enhancement in Decoding MPEG-2 Video</i> , IEEE Transactions on Circuits & Systems for Video Tech., Vol. 12, No. 9, September 2002 (“Martins”) |
| Ex.1008 | U.S. Patent Publication No.2008/0137753 (“He”) |
| Ex.1009 | U.S. Patent No. 5,886,736 (“Chen”) |
| Ex.1010 | Schwarz et al., <i>Overview of the Scalable Video Coding Extension of the H.264/AVC Standard</i> , IEEE Transactions on |

Circuits & Systems for Video Tech., Vol. 17, No. 9, September 2007 (“Schwarz”)

- Ex.1011 U.S. Patent Publication No. 2009/0154567 (“Lei”)
- Ex.1012 Park et al., *Super-Resolution Image Reconstruction: A Technical Overview*, IEEE Signal Processing Magazine, pp. 21-36, May 2003 (“Park”)
- Ex.1013 English Machine Translation of Japanese Patent Publication No. 2008053848 (“Hatanaka”)
- Ex.1014 U.S. Patent Publication No. 2010/0165077 (“Yin”)
- Ex.1015 U.S. Patent Publication No. 2009/0257664 (“Kao”)
- Ex.1016 U.S. Patent Publication No. 2009/0046995 (“K anumuri”)
- Ex.1017 U.S. Pat. 6,470,051 (“Campisano”)
- Ex.1018 U.S. Patent Publication No. 2006/0013306 (“Kim”)
- Ex.1019 Lu et al., Mechanisms of MPEG Stream Synchronization, ACG SIGCOMM Computer Communication Review, Vol. 24, Issue 1, pp. 57-67 (January 1994) (“Lu”)
- Ex.1020 U.S. Patent Publication No. 2003/0021345 (“Brusewitz”)
- Ex.1021 U.S. Patent Publication No. 2003/0021347 (“Lan”)
- Ex.1022 EE|Times, *How Video Compression Works* (Aug. 6, 2007), available at <https://www.eetimes.com/how-video-compression-works> (retrieved Dec. 26, 2023) (“EETimes”)
- Ex.1023 Bier, *Introduction to Video Compression*, Berkeley Design Tech., Inc. (October 2005), available at https://www.bdti.com/MyBDTI/pubs/20051024_GSPx05_Video_Intro.pdf (retrieved Dec. 30, 2023) (“Bier”)
- Ex.1024 U.S. Patent No. 7,379,496 (“Holcomb”)
- Ex.1025 U.S. Patent Publication No. 2008/0043832 (“Barkley”)

- Ex.1026 U.S. Patent No. 6,173,013 (“Suzuki”)
- Ex.1030 Ely, *MPEG Video Coding: A Simple Introduction*, EBU Technical Review (Winter 1995) (“Ely”)
- Ex.1031 Japanese Patent Publication No. 2007316161 A with English Machine Translation (“Phek”)
- Ex.1032 ITU-T Recommendation H.264 (11/07)
- Ex.1033 Karczewicz et al., The SP- and SI-Frames Design for H.264/AVC, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, v.13 (July 2003)
- Ex.1034 ISO/IEC 11172-2: 1993 (E) (MPEG1 Video)
- Ex.1035 ISO/IEC 13818-2: 2000 (E) (MPEG2/H.262 Video)
- Ex.1036 ITU-T Recommendation H.263 (02/98)
- Ex.1037 Reader, Intraframe and Interframe Adaptive Transform Coding, Efficient Transmission of Pictorial Information, SPIE v.66, 108 (1975)
- Ex.1038 Reader, MPEG4: Coding for Content, Interactivity, and Universal Accessibility, Optical Engineering 35(1), 104 (January 1996)
- Ex.1039 Reader, MPEG Patents, in MPEG Video Compression Standard, Chapter 16, at 357-362 (Mitchell et al ed.) (1996)
- Ex.1040 Radha et al., The MPEG-4 Fine-Grained Scalable Video Coding Method for Multimedia Streaming Over IP, IEEE Transactions on Multimedia, v.3, 53 (March 2001)
- Ex.1041 H.264 / MPEG-4 Part 10 White Paper, January 31, 2003 (downloaded from <https://github.com/yistLin/H264-Encoder/blob/1cdaf090b63642932ed726b102ed2cf80908edbb/doc/H.264%20%3A%20MPEG-4%20Part%2010%20White%20Paper.pdf> on Jan. 5, 2024)
- Ex.1042 ISO/IEC 14496-2: 2004 (MPEG4)

Ex.1043	ISO/IEC 13818-2: 1996 (MPEG2)
Ex.1044	EP 1811786 to Wang (“Wang”)
Ex.1046	U.S. Patent Application Publication No. 2007/0140350 (“Sakazume”)
Ex.1048	Loop Filters in the Okubo RM2 (1986)
Ex.1049	U.S. Patent No. 6,580,754 to Wan et al. (“Wan”)

I. Professional Background

A. Summary

4. I am a digital media consultant providing technical, business development and intellectual property consulting services in the areas of digital media including digital imaging, digital video, digital audio and digital speech. Applications include consumer audio and video transmission and storage; video, audio and speech compression; real-time video processing and display; digital speech processing; image/video/audio systems architecture; and image/video/audio chip architecture. I have held this position since 2001 and have consulted for over 80 clients.

B. Education

5. I received my Doctoral degree in 1974 from the University of Sussex, England. My thesis was titled “Orthogonal Transform Coding of Still and Moving Pictures.” The research for my thesis was performed in residence at the Image Processing Institute, University of Southern California, Los Angeles.

6. I received my B. Eng. Degree with Honors in 1970 from the University of Liverpool, England, in the field of electronics.

C. Work Experience

7. From 1970 to 1973, I performed my graduate research in video compression. I was one of the first to perform a type of image coding (adaptive block transform coding) and the first to apply this type of coding to video. This is described in my thesis and summarized in an SPIE paper. Ex.1037. These techniques underlie the audiovisual coding standards known as MPEG (Moving Picture Experts Group), H.26x, and virtually all other video compression schemes today.

8. From 1975 to 1981, I studied, designed and developed systems for military imaging systems including real-time image and video reconnaissance systems and battlefield management systems.

9. In the early 1980s, I taught classes at Santa Clara University, California in digital signal processing and digital image processing.

10. From 1982 to 1989, I architected and led hardware and software engineering teams in the design of systems for real-time imaging for military, medical and earth resources applications. In 1983, I designed an image warping system that could perform geometric warping of images with $1/16^{\text{th}}$ pixel accuracy in near-real-time. The operations included the simple operations in the plane of the

imaging system of translation, spatial scaling – enlargement and minification – and rotation plus higher-order operations such as perspective warping. My team implemented the system in hardware on a single VME board under software control. Significant sales ensued over the next years with a pull-through effect on sales of the complete image processing system in which it was an option. This product design was published in a 1984 paper¹.

11. From 1990 to 2001, I architected and led hardware and software engineering teams in the design of semiconductor chips and systems for real-time imaging in digital consumer audio/video applications, including videoconferencing (with speech coding), broadcast TV and DVD.

D. Standardization

12. In 1990, I became an accredited delegate to the Moving Picture Experts Group – MPEG. I, and my team contributed to the technical work for all three parts of the standard – Systems, Video and Audio – and I participated in the management of the MPEG committee.

13. I was Head of Delegation (HoD) to MPEG for the United States in 1991-1992, during which time the MPEG1 standard was completed and the MPEG2

¹ Adams J, Patton C, Reader C, Zamora D, Hardware for Geometric Warping, Electronic Imaging, April 1984.

standard was successfully positioned as the global standard for digital video broadcasting and recording (DVD). I was the Editor in Chief of the MPEG1 standard. I personally reviewed and edited all three parts of the standard in detail, and wrote much of the informative annex for the MPEG1 standard. I participated in the technical and management work for development of the MPEG2 standard.

14. I chaired the implementation subcommittee that analyzed MPEG1 Audio (Levels I & II aka MUSICAM; Level III aka mp3), Dolby AC3 and other proposed audio compression algorithms including legacy algorithms for complexity and cost of implementation.

15. I was a co-founder of the MPEG4 standard and chaired the subcommittee from inception for 2-1/2 years, beginning in 1993, following which I chaired the MPEG4 Requirements subcommittee for a further 2 years. These subcommittees established many of the fundamental principles of the MPEG4 standard, including object-based coding, software-based implementation, and development of the bitstream as a syntactic language. *See, e.g., Ex.1038.* MPEG4 focused on audiovisual low bitrate coding, including low bitrate speech coding.

16. The MPEG4 standard in ISO and the H.263 standard in ITU-T were developed collaboratively, with the H.263 Rapporteur and I synchronizing meetings and establishing profiles in each standard that are precisely compatible.

17. I was instrumental in establishing the work on Advanced Audio Coding (AAC).

18. I initiated the work in Synthetic-Natural Hybrid Coding (SNHC), and personally contributed to the work on compression of 3D graphics in the area of error resilient coding.

19. In the early 2000s, I was an invited expert to the Joint Video Team (JVT) established by ISO and ITU to develop the H.264 standard (also denoted MPEG4 Pt. 10, AVC). I was also invited to become an officer of the China National Standards committee AVS, in which I chair the Intellectual Property Rights Subgroup.

20. I have closely followed the developments of the H.265 and H.266 standards.

E. Intellectual Property Rights

21. In 1993 I was hired by CableLabs to be the technical expert for establishing the MPEG Patent Pool (Now MPEGLA). In the course of creating a list of essential IP to practice the standard, I reviewed approximately 10,000 abstracts and 1,000 patents. This is summarized in a chapter of the MPEG book by Mitchell et al. Ex.1039.

22. Multiple companies have hired me to assist in developing portfolios of their standards-essential patents.

23. In 2002, I was hired by 10 companies to evaluate the standards essential patent environment for the nascent H.264 standard.

24. In 2003, I was hired to evaluate the standards essential patent environment for the nascent national China AVS standards.

25. In 2013, I was hired to evaluate the standards essential patent environment for the nascent AV1 standard.

26. In 2017, I was hired to evaluate the standards essential patent environment for the H.265/HEVC standard.

27. I am the Co-Director of the China AVS Patent Pool Administration. I lead the negotiations for sub-licensing of the AVS standards. In 2022, I played a leading role in the adoption of the AVS3 video standard by DVB.

28. I have performed expert consulting and expert witness work for patent holders and defendants in patent licensing negotiations and litigation.

F. Curriculum Vitae

29. Additional information concerning my professional publications and presentations in the field of digital video and cases in which I have served as an expert are enumerated in my current Curriculum Vitae, a copy of which is attached as Exhibit A. This Curriculum Vitae lists many publications authored or co-authored by me; and lists the cases in which I have testified via depositions and trials.

II. Digital Video Technologies

30. Below I describe the state of digital video coding as of the '448 Patent's earliest claimed priority date (May 30, 2008), and summarize the development of technologies underlying the features and methods recited in claims 1-3.

A. Analog Video

31. Like film, video comprises a rapid sequence of still images that the human vision system integrates into a perception of motion. The “moving images” are often called video frames and video pictures.

32. All-electronic television² was invented by Philo Farnsworth in the late 1920s. Farnsworth developed the technique of raster scanning images and produced cameras and CRTs (cathode ray tubes) that could synchronously acquire and reproduce moving images.

33. All-electronic television service was introduced by the BBC in England in November 1936. Broadcasting of television programs continued until interrupted by the Second World War. Television service began in the U.S. in July 1941. These services provided only black-and-white video until 1953, when the NTSC standard in the U.S. was revised to include color. The “resolution” of these systems was specified in the number of raster-scan “lines”. Analog broadcasting

² Electro-mechanical television systems were also introduced.

continued in the U.S. for almost seventy years. Full-power analog television broadcasting ceased in 2009 under order from the FCC.

B. Digital Video

34. Digitization of analog signals in general was introduced in the late 1940s with the name “pulse code modulation” (PCM). Fundamental sampling theory showed that a bandwidth-limited analog signal could be sampled at twice the rate of the highest frequency in the band, and still be perfectly reconstructed from such samples. The sample amplitudes could be represented by the value of a number, e.g., a decimal value or a binary value. In the case of image or video data, it has been widespread practice to represent this value by an unsigned 8-bit integer. Analog TV signals that scanned a video frame were digitized into two-dimensional arrays of digital picture elements, known as “pels” or “pixels.” Digital video, like analog video (and film) comprises a rapid sequence of frames.

35. While a correctly bandpass filtered and sampled digital image, such as a video frame, theoretically includes spatial frequencies up to the Nyquist limit, in practice digital images contain spatial frequencies only up to a much lower level. There are many reasons, that may depend on the particular application. For example, the optical assembly that captured the image may limit the optical resolution of the data including lens materials and aberrations. Another example applies in situations where the imaging occurs in an outdoor environment in which atmospheric

turbulence exists between an object being imaged and the image capture system – we may see this ourselves as the heat haze above a highway when driving in hot weather. Another example includes electronic noise in the image capture and processing system. Therefore, while the scene being imaged – the object space – may contain spatial frequencies up to a given level, a digital imaging system designed with a Nyquist frequency at that level will not in practice contain spatial frequencies as high as that. We may refer to such an image as blurred.

36. In this context, it is important to understand what is meant by the term “resolution”. A distinction should be made between the resolution of the information in an image (such as a video frame) and the data representing the information in the image. The latter comprises the array of pixels, referred to above, and it is common in everyday speech and within the imaging/video community itself to refer to the number of pixels or the number of lines as the “resolution” of the image. The term “spatial resolution” is also frequently used in the art. As just explained, typical digital images do not contain spatial frequencies up to the theoretical limit given by the number of pixels/lines. For example, many high-definition television (HDTV) receivers comprise a pixel array that vertically has 1080 lines. There are two principle broadcast formats – NBC, CBS and PBS broadcast in a “1080i” format (where the i denotes interlace-scan), while ABC and Fox broadcast in a “720p” format (where the p denotes progressive format). Both

are presented on the HDTV receiver in a 1080p format by processing the 1080i broadcasts to deinterlace the signals, and upsampling the 720p broadcasts to 1080 lines. But according to the Kell principle, the 1080i signal only contains about 700 lines of perceivable information and when either format is converted to 1080p format, neither of them has 1080 lines of “information” even though they both have 1080 lines of “resolution”.

37. PCM was followed in 1952 by the invention at Bell Labs of differential pulse code modulation, (DPCM), which was immediately applied to video. Because the correlation between pixels in typical scenes was high, it was efficient to code the pixels by successively predicting each pixel from the preceding pixel or pixels. AT&T demonstrated its Picturephone in 1964 at the World’s Fair. Subsequently, commercial Picturephone service was introduced in 1970, using DPCM coding.

38. Digital video was introduced in the professional television studio environment with the CCIR 601 standard that was developed in the early 1980s and published in 1982 (now known as the ITU-R BT.601 standard). The standard covered PCM data formats for the NTSC and PAL video standards and facilitated the development of modern production and post-production studios, involving non-linear editing, digital switchers, and mixers, as well as digital special effects.

39. Digital video in its raw format is highly redundant. Each video frame comprises a regular, rectangular array of pixels. The density of the pixel array is

chosen such as to portray the highest spatial frequencies enabled by a particular video system. Historically, the earliest systems portrayed 400-500 lines of resolution. The so-called HDTV systems introduced in the 1990s doubled the number of lines to 1000 lines, and the recent so-called 4K systems double the number of lines again to 2000 lines. The imminent next generation of so-called 8K systems will further double the number of lines to 4000 lines, which will approximate the resolution of a 35mm film system.

40. Historically the frame rate was 30 frames per second, but one of the HDTV formats doubled this to 60 frames per second.

41. Transmission or storage of raw digital video data requires huge volumes of data and very fast data rates. Transmitting standard definition color TV with a format of 720x486 pixels and 30 frames/s would require over 250Mbits/s. HDTV would require approximately 5 times as much data and so-called 4K TV (UHDTV) at 60 frames/s would require approximately 40 times as much data.

42. Within a given scene being imaged by a video system, the content is very likely not to contain the highest level of detail except in a few, small areas. While a prescribed density of pixels may accurately describe such high detail areas, that density of pixels is overkill for the rest of the scene to a greater or lesser degree. So the typical digital video frame is spatially redundant. For example, if the video

scene contains the sky, or a uniformly painted wall, adjacent pixels will have almost identical intensity and color.

43. The frame rate of the video is established to smoothly portray the fastest motion the system is designed to support. But such motion may occur only a fraction of the time, and may occur only over a portion of a frame spatially. So the typical video sequence of frames is temporally redundant. For example, sportscasters sitting at a desk demonstrate little motion over only a portion of the scene, but when they cut away to basketball, the action is frenetic and may be scene-wide. In the former case the pixel values will be almost identical from one frame to the next over most of the frame.

C. Image Restoration; Superresolution

44. The field of image restoration is concerned with providing the maximum image quality possible within a pixel array of a given dimension (“pixel resolution”). The raw image may be degraded by various aberrations and corrupted by noise, as discussed above. The fine details within the image – which in general correspond to the high frequency content of the image may not be visible to a viewer. Image restoration is the class of processing that reverses the aberrations and suppresses the noise, thus exposing the fine details.

45. Research in this field is over 50 years old, and comprises multiple approaches, involving linear and non-linear filtering, statistical methods and

perceptual models. Early work in this field is compiled in the book Digital Image Restoration, Prentice-Hall, 1977 by Andrews and Hunt. Approaches to the problem include mathematical modeling of the aberrations and applying the inverse of the model to correct for the aberrations, or statistical modeling of the noise and applying digital signal processing to filters out the noise.

46. One member of the class of image restoration techniques is superresolution. The term implies a process to create an output image having spatial frequencies higher than the theoretical limit given by the dimensions of the image pixel array of the input image. This is not possible and practically the term is used to provide enhanced spatial frequency content in two different ways. In the first way, the term refers to restoring the spatial frequency content of an image within the theoretical limits of the pixel array dimensions. This means that the information content of the superresolution processed image, i.e., the “information resolution” is higher than that of the input image, within the same pixel array dimensions, i.e., the “pixel resolution”. However, in the other way, the term is used to describe a process in which the processed image possesses a pixel array having larger dimensions than the input image. For example, the output superresolution image may have twice as many pixels horizontally and vertically (four times as many pixels in total), i.e., higher “pixel resolution”.

47. Merely doubling the number of pixels horizontally and the number of pixels or lines vertically in the representation of an image does not increase the resolution of the image³. Using superresolution processing to produce an output image that has double the pixels in each dimension may increase the information content (range of spatial frequencies) of the image, i.e., the true information “resolution” of the image, beyond that supported by the original pixel dimensions and at the same time the output image has twice the “resolution” in pixels horizontally and vertically. But the two are not synonymous.

48. Improving the spatial frequency content of an image to provide “superresolution” when the raw image spatial frequencies are limited by factors such as physical phenomena in the image capture process may seem impossible. But if multiple images of the same scene (object space) are available, and these images are not aligned with each other by translational differences that are fractions of the dimensions of a single pixel, then the assembly of such images does contain higher

³ One can easily check this on one’s own computer – zooming in (scaling up) an image that is initially being displayed with each pixel in the image corresponding to a pixel on the display will not cause the displayed image to become sharper, i.e., show more information, instead it will look blurry and eventually the individual pixels will become visible as square blocks on the screen.

spatial frequency content than any one member of the assembly. A very well-established superresolution method in the art is to register such multiple images using sub-pixel accuracy, and filtering the registered image to output an image with higher spatial frequencies than any one of the input images.

49. Many prior art publications describe this approach to superresolution, including admitted prior art in the '448 Patent. See “Super-Resolution Image Reconstruction, by Sung C. P.⁴ and Min K. P.: A Technical Overview, IEEE Signal Proc. Magazine, Vol. 26, No. 3, pp. 21-36, 2003”, Ex.1001 ['448 Patent], 22:13-16. The source of the multiple images of the same scene that are to be registered may be satellite images taken on multiple passes over the same location. This is described in a 1984 paper by Tsai: Tsai R Y, Huang T S, Multiframe image restoration and registration, Advances in Computer Vision and Image Processing, Vol. 1, T. S. Huang, ed., Jai Press, pp. 317-319, 1984. The '448 Patent, for example, refers to the Park prior art reference (Ex.1012 (Park, *Super-Resolution Image Reconstruction*)) in discussing super-resolution enlargement. Ex.1001 ['448 Patent], 22:13-16. Park summarizes super-resolution processing as follows:

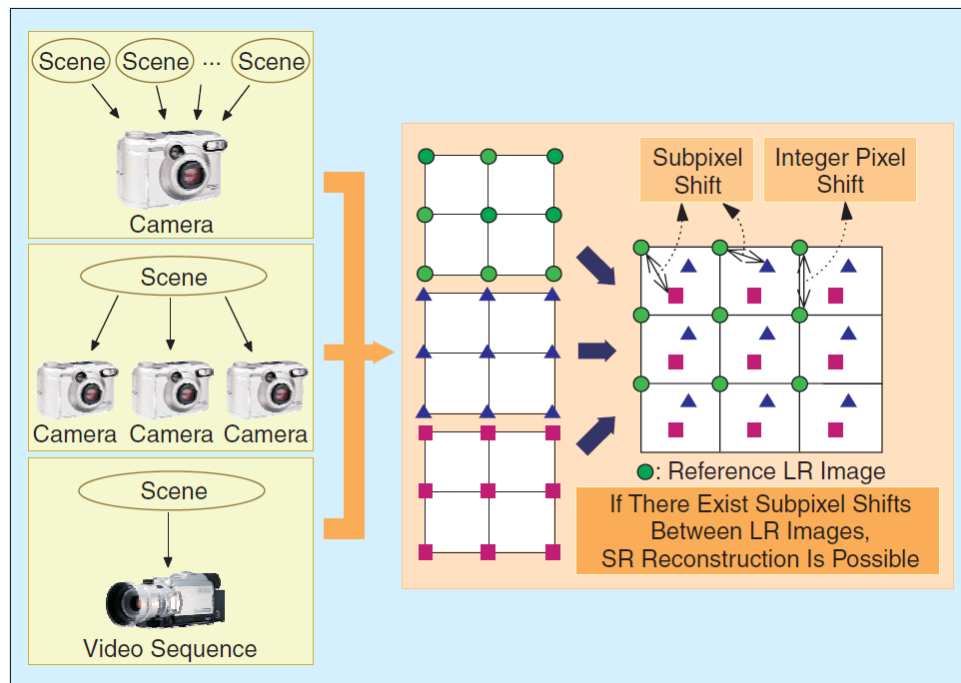
“How can we obtain [a high-resolution (HR)] image from multiple [low-resolution (LR)] images? The basic premise for increasing the spatial resolution in [super-resolution (SR)] techniques is the

⁴ The patent errs in misunderstanding the order of Asian surnames and given names. The surnames are Park and the initials are C. P. and M. K.

availability of multiple LR images captured from the same scene In SR, typically, the LR images represent different ‘looks’ at the same scene. That is, LR images are subsampled (aliased) as well as shifted with subpixel precision. ... If the LR images have different subpixel shifts from each other and if aliasing is present, ... then ... the new information contained in each LR image can be exploited to obtain an HR image.”

Ex.1012 [Park], 2.

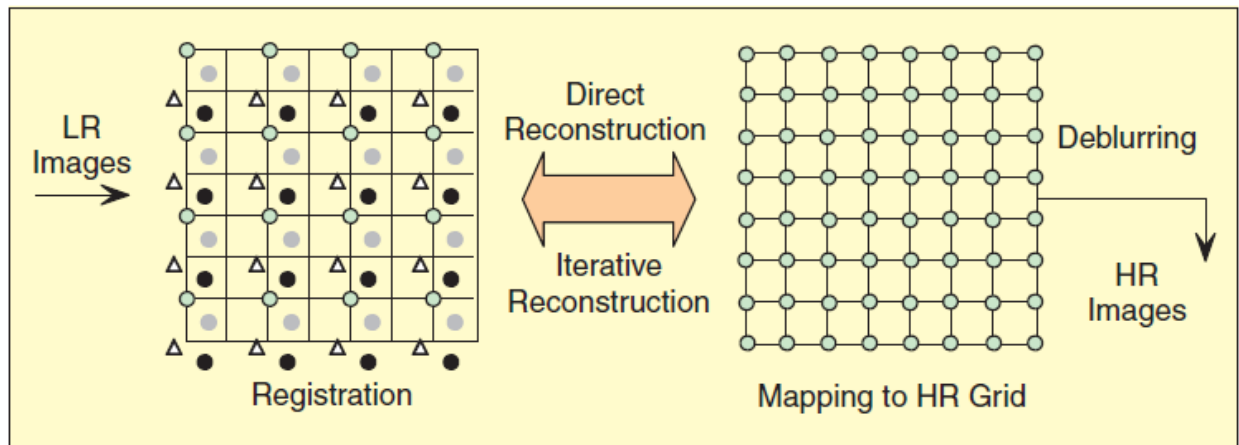
50. Park’s FIG. 1, for example, illustrates how three low-resolution pictures of the same scene, captured individually by several still cameras or sequentially by a video camera, naturally contain pixels that are spatially offset from each other by amounts smaller than the size of individual pixels in the low-resolution grid (i.e., subpixel shifts):



▲ 1. Basic premise for super resolution.

Ex.1012 [Park], FIG. 1

51. According to Park, typical super-resolution processing involves registering pixel locations from several low-resolution pictures, mapping pixels from the low-resolution pictures to a high-resolution grid, interpolating any missing pixels in the high-resolution grid, and performing deblurring and denoising operations. Ex.1012 [Park], 5. Park's FIG. 7 illustrates these concepts:

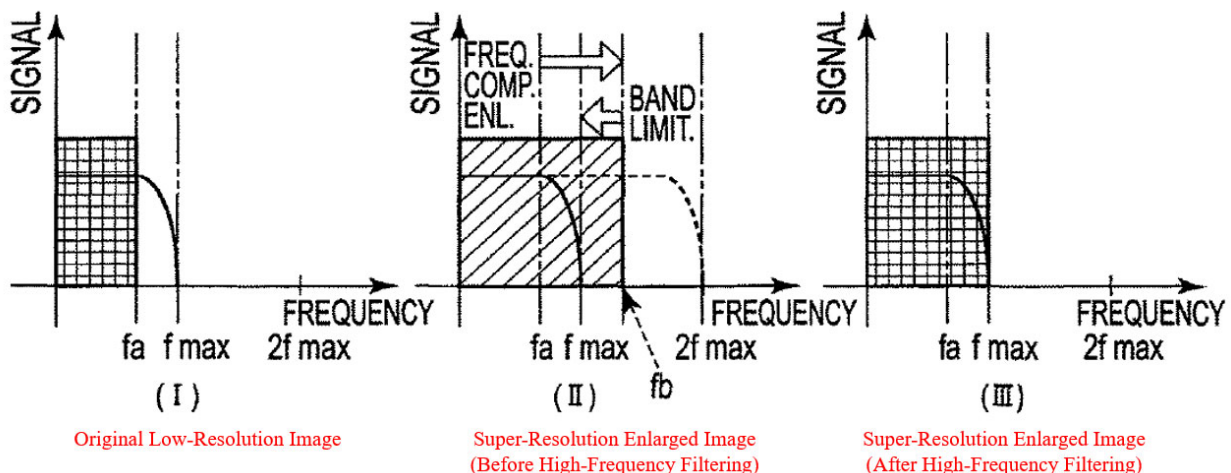


▲ 7. Registration-interpolation-based reconstruction.

52. Another source is the frames of a video sequence. Video standards are designed to capture and display video images that are close in time so there is a limited amount of motion between each pair of images and the human vision system integrates the motion from one image to the next to create the illusion of continuous motion. Most conveniently, the sub-pixel registration of such frames is facilitated by the process of motion-compensated prediction used in video coding standards to code a sequence of such frames. The motion vectors of the coded sequence already provide the data for registering the frames for superresolution processing.

53. Techniques for reducing the pixel resolution of a super-resolution enlarged picture to maintain the enhanced quality of such a picture while supporting displays or other processes at a lower pixel resolution were well-established before the '448 Patent. Park, for example, explains that “[a] related problem to [super-resolution (SR)] techniques is image restoration” where “the goal ... is to recover a degraded (e.g., blurred, noisy) image, but it does not change the size of image.” Ex.1012 [Park], 2. Martins further describes techniques for “decimating” super-resolution enlarged pictures in a video sequence to obtain “higher fidelity” video in a desired lower-resolution format. Ex.1007 [Martins], § IV.

54. FIG. 4(c) of the '448 Patent illustrates how superresolution increases the spatial frequency content of an image from a lower frequency “ f_a ” to a higher frequency “ f_b ” (which may be subsequently filtered to a desired frequency “ $f_{\max}$ ”):



Ex.1001 ['448 Patent], FIG. 4(c) (annotated)

D. Image Warping

55. An image or video sequence may be captured of a scene and an application may require the view of the scene to be oriented differently. In the simplest case, the image must be translated in the plane of the imaging system by some number of pixels horizontally and/or vertically. The translation may be needed with fractional pixel (sub-pixel) accuracy. Another case is one in which the scale of the captured image must be changed, i.e., it is desired to zoom in (enlargement) or zoom out (minification) on the image. A further case is rotating the image and beyond that more sophisticated warping operations may be needed such as perspective view changes.

56. To implement image warping two operations are required. First the addresses of the output pixels must be calculated in input image coordinate space with fractional pixel (sub-pixel) accuracy. Then the output pixel values must be computed from the input pixel values by interpolation from the input pixels in the neighborhood of each output pixel based on the fractional pixel spacing.

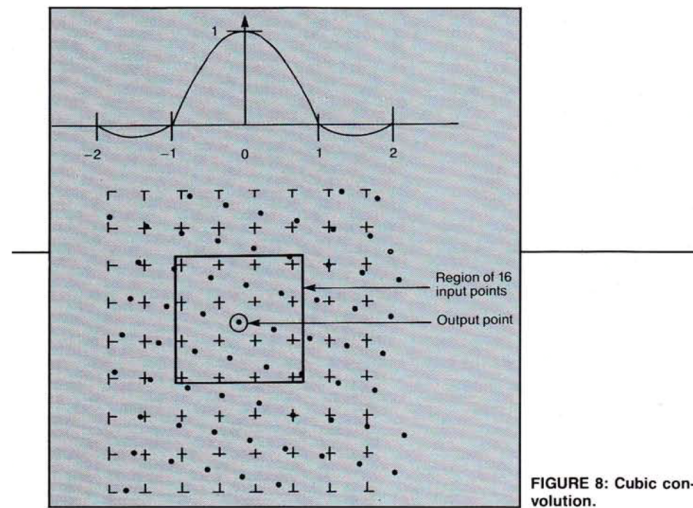
57. In the image warper I designed 40 years ago, the address calculations were performed using a polynomial equation of the appropriate order. The interpolations could have been performed by first interpolating the entire input image to create values for sub-pixels at every sub-pixel location. That would have constituted an “enlargement” of the input image. Then the desired pixels for the

output image could have been selected from the enlarged array using the computed output addresses in the input coordinate space. But that would have required computing and storing the enlarged image, which would have been larger by the square of the sub-pixel accuracy (256 times larger given the $1/16^{\text{th}}$ pixel accuracy of my design), while only a small fraction of the values would have been selected for the output image. Such an “enlargement” procedure would have been incredibly wasteful. Instead, for each pixel in turn, the address of each output pixel in input space coordinates was used to define a region of the input pixels matching the size of the interpolation kernel (4x4 pixels) and the output pixel was interpolated from that region.

58. My 1984 paper illustrates interpolation of output pixels using the cubic spline interpolator. It would have been extremely wasteful to “enlarge” the input image by a factor of sixteen in both dimensions to provide the $1/16^{\text{th}}$ pixel accuracy of the design. So the value of each output pixel was calculated by superimposing the contributions from a 4x4 array of neighboring input pixels. This approach enabled direct calculation for each output pixel, avoiding both “enlargement” of the input image and “resolution conversion”, i.e., filtering and downsampling of the output image.

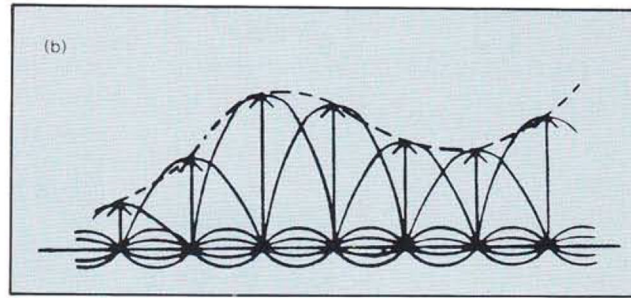
59. Note this technique enabled interpolation of output pixels from input pixels with sub-pixel accuracy for a wide range of geometric warping relationships,

including simple translations in plane, simple enlargement or shrinking, simple rotation in plane, or advanced warping such as perspective mapping.



60. Figure 8 illustrates output of a rotated image (dot pixels) from an input image (cross pixels). The interpolation filter kernel is illustrated (in one dimension) at the top, and was stored in a lookup table with unit peak value and $1/16^{\text{th}}$ pixel accuracy⁵. The 4x4 region of neighborhood pixels is illustrated below, and for each output pixel, the value of each input pixel was used to weight the filter kernel at the corresponding fractional distance between the input and output pixels. An example of such superposition of weighted kernels to interpolate interstitial values is shown as the dotted line in Figure 7b (in one dimension).

⁵ The kernel size was thus 64x64, but by exploiting two-dimensional symmetry, only one quarter of the kernel needed to be stored in the lookup table.



61. The patent specification discloses use of spline functions for interpolation. *See, e.g.,* Ex.1001 ['448 Patent], 25:5-35, 27:10-17.

E. Video Coding

62. Video coding, also known as video compression, is the art of removing the spatial and temporal redundancy to enable storage or transmission of digital video data with vastly fewer bits, while preserving the quality of the raw video data to the maximum extent possible. Expressed another way, video coding is the art of representing the information in a video sequence with the least data (i.e., the fewest bits) and the highest reconstructed quality.

63. Video coding is not performed by a single algorithm, rather it comprises a set of video coding tools, each designed to tackle a particular aspect of reducing the redundancy of raw video data. Some video tools tackle the spatial redundancy, while other tools tackle the temporal redundancy. In each case, the tools attempt to identify and extract the visual information perceived by the human vision system. This information is then encoded into a video bitstream.

64. The fundamental research for video coding was performed in the 1960s and 1970s. By the end of the 1970s, a so-called hybrid architecture had been developed that remains today as the foundation for almost all video coding standards.

65. Spatial redundancy is reduced using adaptive block-based transform coding. In the spatial domain, each individual pixel has the same information content, i.e., entropy. But since in a natural video scene, many areas contain little or no detail, i.e. only low spatial frequency information, adjacent pixels are highly correlated and spatially transforming the pixel data results in transform coefficients that only have significant magnitude for low spatial frequencies while a significant majority of coefficients have negligible or zero magnitude. The information is then represented by a small number of coefficients, which may be coded efficiently. This technique is efficiently implemented using small blocks.

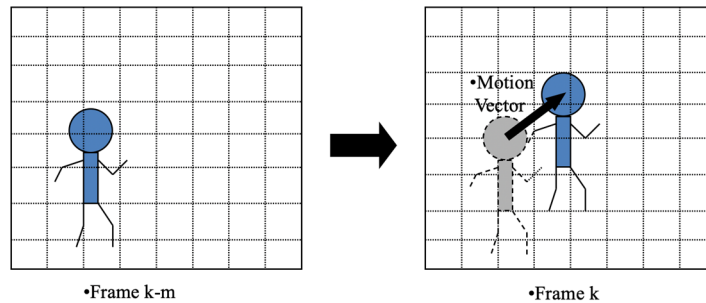
66. Temporal redundancy is removed using block-based motion compensation, first invented in 1974⁶. This is based on a model of motion in which objects within a scene generally do not change, but rather move from one location

⁶ Taki Y, Hadori M, Tanaka S, Inter-frame encoding for motion following, 1974 Natl. Conf. of the Inst. of Electronics, Information, and Communications Engineers, July 1974 (Translation)

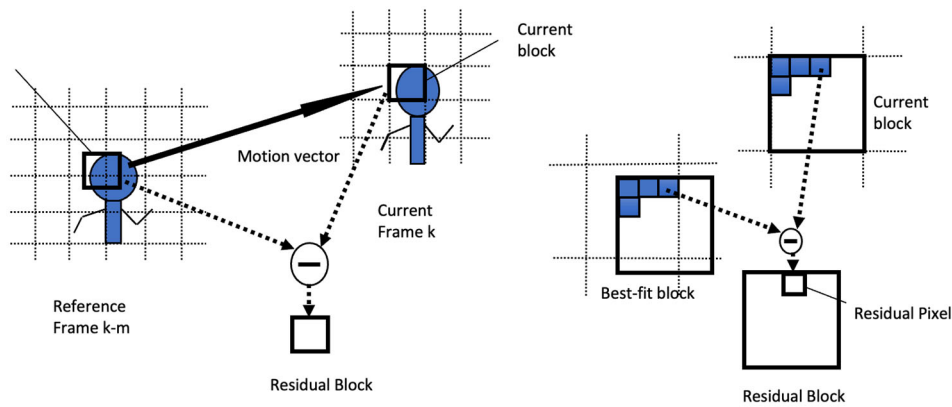
Brofferio S, Cafforio C, Del Re P, Quaglia G, Racciu A, Rocca F, Redundancy Reduction of Video Signals Using Movement Compensation, *Alta Frequenza*, pp. 836-843, Oct 1974

in the scene to another. On a block basis, this model works well. The following figures I assembled illustrate the point:

- **Model: Motion consists of translation in the viewing plane.**



67. Thus for each block in a current frame, a search is made to find a best match to the block in a “reference” frame that has already been coded. The best-matched block is used as a prediction for the current block. The difference between the pixel values of the current block and the pixel values of the prediction block is called the residual data. The location of the best match block in the reference frame relative to the location of the current block is called a motion vector. The residual data block and the motion vector are coded, and typically comprise a small fraction of the amount of data needed to code the current block alone. The following figures I assembled illustrate the point:



68. Research in the early 1980s showed that performing motion compensated prediction with sub-pixel accuracy resulted in significantly better results. The pixel array of the reference image is interpolated to generate sub pixels at fractional locations between the original pixels, for example, $\frac{1}{2}$ or $\frac{1}{4}$ locations. This will result in an enlarged pixel array⁷. Motion estimation is performed for a current block using the interpolated reference image data to find the location with the best match between the current block pixel values and the sub-pixel interpolated reference pixel values, i.e., the pixels of the current block are registered with the interpolated sub-pixels of the reference block. The difference between the current block pixel values and the reference picture sub-pixel values is taken, constituting

⁷ In practice, a preliminary match is found using the reference pixels at integer locations and then “enlargement” is performed of the local region to provide sub-pixel accuracy and a refined match is found.

the residual block data and the motion vector is expressed with both an integer and fractional part.

69. For blocks that cannot be temporally predicted efficiently, spatial prediction within the same frame is applied. In both cases the blocks of residual data may still contain spatial redundancy within the block. This redundancy is reduced using transform coding techniques, as discussed above.

70. The nature of video data is highly variable, both spatially and temporally. Some video sequences contain high levels of spatial detail. Such detail may span across the whole video frame or may be confined to local areas of the video frame. Some video sequences contain high degrees of movement that also may span across the whole video frame, such as when a camera is panning rapidly across a scene, or may be localized within the video frame. Video coding is performed in a highly adaptive way, by breaking each video frame into small blocks and adapting the spatial and temporal coding to the characteristics of the data on a local basis.

71. The temporal and spatial residual information is entropy coded to efficiently represent it and is then formatted into a bitstream.

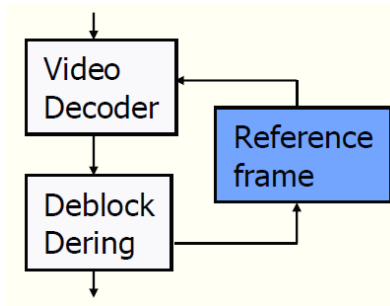
72. The adaptive coding of video results in a highly variable instantaneous bitrate – it is obvious the volume of data produced when a ball is snapped in a football game is vastly higher than that produced during a time-out. As a result, the

entropy coded bitstream at the encoder is buffered in order to smooth out at least some of the variation in data rate before transmission or storage and there is a corresponding buffer at the input to the decoder.

73. Video coding in particular has been characterized by the difficulty and cost involved in implementing coding algorithms for use in mass-market consumer applications. The volume of video data is very large, and the requirement for presenting video data that has been transmitted or retrieved from storage in real-time has been very challenging. As a consequence, while researchers in the field have conceived of sophisticated mathematical algorithms and procedures for video coding for decades, practical implementation has had to wait for semiconductor technology to pass certain thresholds.

F. Loop Filters

74. Many conventional video coding systems before the '448 Patent included “loop filters” that were designed to enhance reference pictures used in motion compensation. The following figure illustrates the concept of loop filtering in the context of “deblocking” or “deringing” filters that enhance reference pictures used for motion compensation:



Ex.1023, Slide 34

75. As shown above, the “deblocking” or “deringing” loop filter enhances a decoded (reconstructed) picture from the video decoder before that picture is used as a reference for the prediction of pixels in other pictures of the video sequence. Ex.1023, Slide 34.

76. Deblocking and deringing filters are two common types of loop filters that reduce artifacts between adjacent blocks of a reference picture that have been predicted from disparate blocks in one or more reference frames or intra-predicted from other blocks in the same picture. The enhanced quality of a picture before deblocking filtering (left) and after deblocking filtering (right) is shown below, as an example:



Ex.1023, Slide 34

77. Deblocking and deringing filters were common, but were not the only type of loop filters known before the '448 Patent. Many types of loop filters were in fact known before the '448 Patent, each of which functions to enhance the quality of a reference picture in a coding loop. *See, e.g.*, Ex.1008 [He], [0006]; Ex.1015 [Kao], [0068]; Ex.1011 [Lei], [0022]-[0023], FIGS. 2-5; Ex.1048 [Doc144], p. 1.

78. While loop filters can improve decoding quality and compression, their benefits must also be weighed against the cost of additional processing load incurred by the filtering. *See, e.g.*, Ex.1023 [Bier, *Introduction to Video Compression*], Slides 34-35; Ex.1008 [He]; *infra*, §VI.D. For this reason, prior art video coding algorithms implemented selective loop filtering. Selective loop filtering allows the coder to decide whether to apply loop filtering to a reference picture based on factors that balance the marginal image quality improvement realized by the loop filter, the amount of additional compression realized by the loop filter, and the cost of loop

filtering in terms of increased computational load and power consumption. *See, e.g.*, Ex.1008 [He], [0005]-[0008], [0028].

G. Video Coding Standards

1. Background

79. Two international bodies are responsible for audiovisual standards – the ITU-T and ISO/IEC. Throughout the last three decades, these bodies have collaborated to develop generations of standards, most frequently publishing identical texts under their own standardization programs.

80. In general, the standardization process has followed the practice of beginning with a requirements gathering stage, leading to a call for proposals and evaluation stage, and then a development stage. In the development stage, the draft standard, and complementary reference software, are iterated from one meeting to the next.

81. Owing to the highly variable nature of video data, (and non-stationary statistics) the experts propose techniques to develop a particular coding tool and perform experiments using the reference software and selected test sequences to determine whether or not, and under what circumstances, a proposed technique improves the draft coding standard. Frequently, a technique brings improvement under specific, limited circumstances, so it is common practice for a tool to be switched on and off for local regions of the video sequence (spatially and/or

temporally), or for parameters of a coding tool to be locally adjusted. In this sense, the development of a standard and attendant adoption of coding tools is an ad hoc process based upon empirical evidence.

82. In the final standard, this appears in the syntax specification of the coded bitstream as a combination of flags and sections of optional syntax selected by the state of the flags; parameters that adjust/adapt video coding tools; and indices that specify selections among multiple choices. For example, for motion compensation of a video frame or blocks within a video frame, a selection may be made between prediction from one of a number of reference frames. The encoder determines which reference frame is optimal at any given point and signals the index of that reference frame to the decoder. There are no hard and fast rules for the selection – a reference frame that works best for one block in the current frame may not be the best choice for the next block.

83. The structure of standards for audiovisual coding has comprised an audio/speech coding part, video coding part, and a “systems” part. The audio/speech and video parts establish standards normatively specifying precise formats for coded bitstreams dedicated exclusively to audio/speech and video respectively. These may be called elementary bitstreams. For audiovisual applications such as broadcast television, these elementary bitstreams are combined into a single bitstream by the system part.

2. The H.261 Standard

84. The first video coding standard, directed toward videoconferencing was technically completed in November 1989⁸, and published in 1993⁹. The standard was based on the hybrid motion compensated prediction / adaptive block transform hybrid architecture and included a loop filter that could be switched on and off to provide motion compensated prediction with and without filtering. This is illustrated in a figure of the standard, in which the loop filter F can be seen, together with a signal f that the encoder generates to indicate whether the filter is ON or OFF, and which is multiplexed and transmitted to the receiver. Figure 3/H.261, Document #584, p. 6:

⁸ CCITT SGXV, Working Party XV/1, Specialists Group on Coding for Visual Telephony, Tokyo, 7-10 November 1989, SOURCE: SPECIALISTS GROUP ON CODING FOR VISUAL TELEPHONY, TITLE : DRAFT REVISION OF RECOMMENDATION H.261, Document #584, Nov. 10. 1989.

⁹ ITU-T H.261, TELECOMMUNICATION (03/93) STANDARDIZATION SECTOR OF ITU, LINE TRANSMISSION OF NON-TELEPHONE SIGNALS VIDEO CODEC FOR AUDIOVISUAL SERVICES AT p x 64 kbits, ITU-T Recommendation H.261. (Previously "CCITT Recommendation")

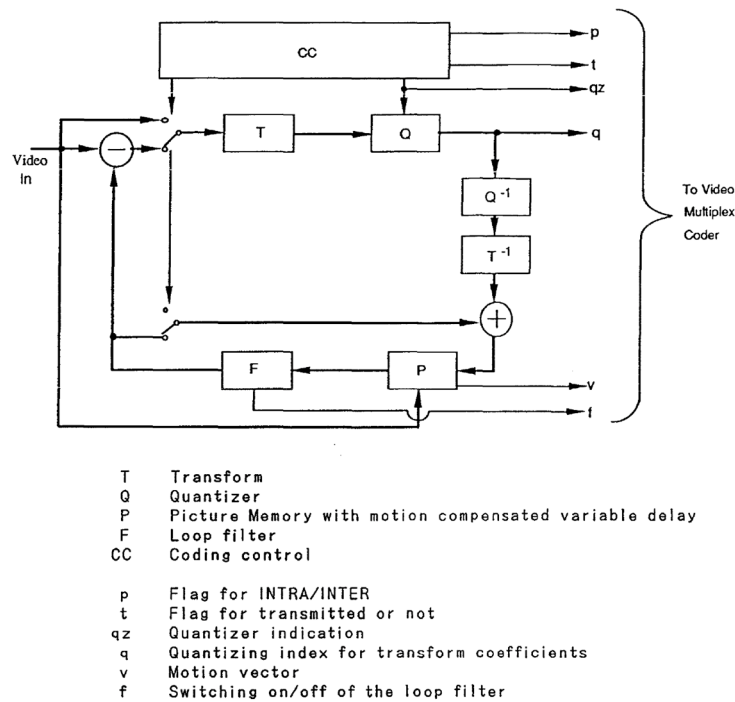


Figure 3/H.261 Source coder

85. One overview of the video coding algorithm is provided, Document #584, p. 5:

“3.2 Video source coding algorithm

The source coder is shown in generalised form in Figure 3. The main elements are prediction, block transformation, and quantisation.

The prediction error (<INTER mode) or the input picture (INTRA mode) is subdivided into 8 pel by 8 line blocks which are segmented as transmitted or non-transmitted. Further, four luminance blocks and the two spatially corresponding colour difference blocks are combined to form a macroblock as shown in Figure 10 of §4.2.4.

The criteria for choice of mode and transmitting a block are not subject to recommendation and may be varied dynamically as part of the coding control strategy. Transmitted blocks are transformed and resulting coefficients are quantised and variable length coded.”

“3.2.1 Prediction

The prediction is inter-picture and may be augmented by motion compensation (§ 3.2.2) and a spatial filter (§ 3.2.3).

86. The loop filter is specified in the standard, Document #584, p. 6:

3.2.3 Loop filter

The prediction process may be modified by a two-dimensional spatial filter which operates on pels within a predicted 8 by 8 block.

The filter is separable into one dimensional horizontal and vertical functions. Both are non-recursive with coefficients of $1/4$, $1/2$, $1/4$ except at block edges where one of the taps would fall outside the block. In such cases the 1-D filter is changed to have coefficients of 0, 1, 0. Full arithmetic precision is retained with rounding to 8 bit integer values at the 2-D filter output. Values whose fractional part is one half are rounded up.

The filter is switched on/off for all 6 blocks in a macroblock according to the macroblock type. (See §4.2.3 MTYPE).”

87. The loop filter in H.261 served two functions – first it provided a degree of smoothing for block edge artifacts, second it provided for more efficient motion estimation and in consequence more coding efficiency. The original proposal for a loop filter was document #144 in November 1986¹⁰, which is incorporated in this

¹⁰ CCITT SGXV, Working Party XV/1, Specialists Group on Coding for Visual Telephony, SOURCE: UK, NETHERLANDS, SWEDEN, TITLE : LOOP FILTERS IN THE OKUBO RM2, Document #144, November 1986.

declaration by reference in full.. The document states the following advantages, Doc. #144, p. 1 (Ex.1048):

- A reduction of high frequency artifacts introduced by MC.
- A reduction in quantisation noise circulating in the feedback loop.
- A reduction in unwanted high frequencies circulating in the feedback loop.
- For systems with motion compensation the filter could be made adaptive as a function of motion vectors with no additional overhead cost.

88. Assuming that the filtered block is coded in an INTER mode, then the high frequency components will actually be coded INTRA and the low frequency components INTER

89. In the conclusion, Ex.1048, p. 4, the document states:

It is clear that the application of a filter within the coding loop as described above improves the quality of the coded pictures. ... However it has been shown that the filter must be capable of being switched on or off ...

3. The MPEG1 Standard

90. The MPEG committee established requirements for the MPEG1 standard and called for proposals in 1989. Many of the resulting proposals were based on the preceding work for the H.261 videoconferencing standard, indeed, MPEG1 may be viewed as an extension of that standard. But for MPEG1, the requirements were focused on a digital replacement for the analog VCR, which did

not have the low-delay requirement of videoconferencing and included trick play modes such fast-forward, fast rewind and reverse play. Multiple responses to the call for proposals proposed groups of pictures as a mechanism to meet these requirements. (For example, reverse play could be performed by decoding the group of pictures in normal, forward play, buffering the decoded pictures and displaying them in reverse order). During the ensuing development of the standard, the Group Of Pictures (GOP) data structure was defined.

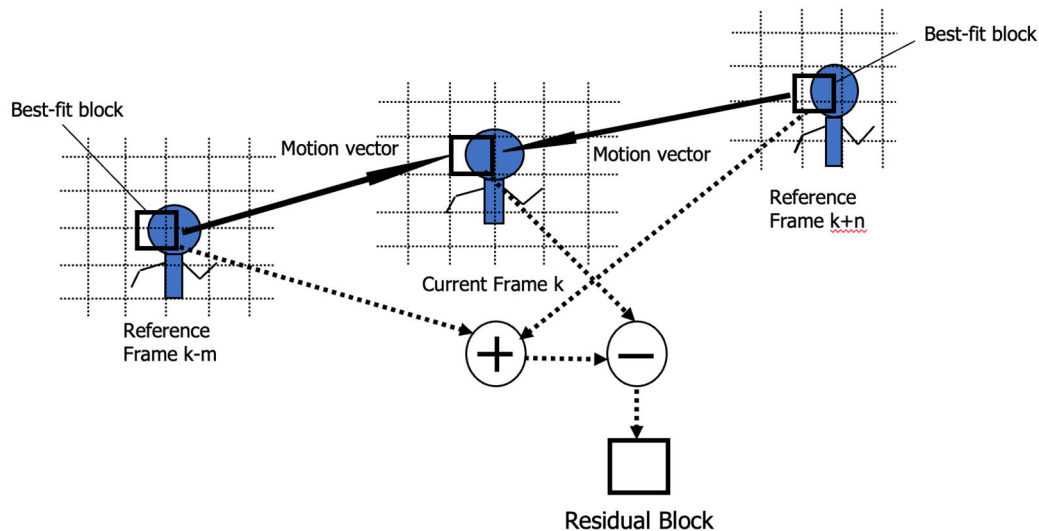
2.1.71 group of pictures [video]: A series of one or more coded pictures intended to assist random access. The group of pictures is one of the layers in the coding syntax defined in this part of ISO/IEC 11172.

Ex.1034 [MPEG1 Video], 6 (original pagination).

91. The group of pictures must contain an I-picture, may contain P-pictures and B-pictures and can be any length greater than or equal to one I-picture. Ex.1034 [MPEG1 Video], 15-16 (original pagination). In MPEG1, the coding order and display order of the pictures of the GOP can be different from each other. *Id.* In trick modes such as reverse play, successive groups of pictures may be decoded in forward order, buffered and then displayed in reverse order.

92. The MPEG1 standard introduced a highly efficient method of temporal coding called bi-directional prediction. This was subsequently adopted by the MPEG2 standard. In this method, multiple video frames are buffered prior to encoding and the coded data for multiple frames are buffered before decoding.

Temporal coding is used first to predict and code the blocks in a future frame that typically may be two or three frames ahead in the sequence of buffered frames. This is called a “P-frame.” Then the blocks in an intervening frame are coded by searching for matching blocks in a previous frame and a future frame. This is called a “B-frame.” The corresponding pixel values in the best match blocks are averaged and the averaged block is used to predict the block in the intervening frame. The resultant residual block and the two corresponding motion vectors are coded.



93. In the MPEG1 standard, motion compensation can be performed with $\frac{1}{2}$ -pixel accuracy.

94. The prediction of P-frames by successive forward prediction and B-frames by prediction from adjacent I-frames and P-frames is illustrated in Figure D.2 of the MPEG1 Standard, shown below. It can be seen that intra-coded frame I2 is not predicted from another frame; P-frames P5, P8 and P11 are forward-predicted

from I-frames or other P-frames, and pairs of frames I2, P5, P8 and P11 are used to predict the corresponding interstitial B-frames.

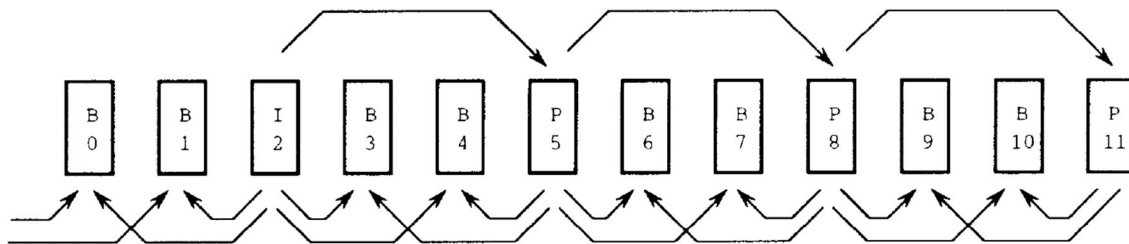


Figure D.2 -- Dependency relationship between I, B, and P-pictures

Figure D.2 illustrates a number of pictures in display order. The arrows show the dependency relationship of the predictive and bidirectionally predictive coded pictures.

Ex.1034 [MPEG1 Video], 54 (original pagination).

4. The MPEG2 Standard

95. The MPEG committee established requirements for the MPEG2 standard and called for proposals in 1991. The requirements were focused first on digital television broadcasting and recording of standard definition video (e.g., NTSC, PAL) at full resolution. As such, MPEG2 comprises an extension of MPEG1 to include coding of interlaced video. The MPEG2 standard, which is incorporated in this declaration by reference in full, also includes optional coding tools for scalable video coding. *See, e.g.,* Ex.1043 [MPEG2]; Ex.1035 [MPEG2/H.262 Video].

96. The MPEG2 standard, officially ISO/IEC 13818-2 was technically completed in 1994, and rapidly became massively successful in the global

marketplace as the standard for digital television broadcasting and digital video recording – DVD. *See, e.g.*, Ex.1043 [MPEG2]; Ex.1035 [MPEG2/H.262 Video].

5. MPEG Scalability

97. In the context of video coding, scalability is a functionality in which a single coded bitstream comprises a hierarchical set of sub-streams representing the video data. The bitstream is hierarchical because a base layer is coded at a given reproducible (decoded, reconstructed) quality level, decoded, and then used to predict a layer that is coded at a higher reproducible (decoded, reconstructed) quality level. This process may be iteratively continued to predict further higher quality layers. The concept is the enhancement layer frames are highly correlated with the base layer frames, so coding only the difference between them requires few bits. It should be noted the decoded lower layers will be buffered and the prediction of higher layers synchronized on a frame-by-frame basis.

98. The complete set represents the video data at the highest quality for reproduction and the highest bitrate for storage and transmission. Subsets of the streams may be stored and transmitted at lower bitrates, or subsets only may be decoded at lower computational cost, leading to lower reproduced quality.

99. A subset may represent the video content at a lower quality, or at a lower spatial resolution, i.e., smaller pixel array, or a subset may represent the video content at a lower temporal resolution, i.e., a lower frame rate.

100. In the MPEG standards, B-frames inherently support temporal scalability because they are not used to predict other frames and thus may not be transmitted, or decoded without affecting the transmission, decoding and presentation of the P- and I-frames. Examples of practical applications for such temporal scalability include a heterogeneous network in which a node on a broadband network interfaces to a narrow band network, at which a full resolution, full bandwidth video bitstream may be forwarded by stripping out the bits used to code the B-frames. Another example is a mobile receiver having limited computational power, that decodes and presents only the I-frames and P-frames.

101. The MPEG2 standard supports three types of scalability – SNR scalability (i.e. quality scalability), spatial scalability and temporal scalability. Ex.1035 [MPEG2/H.262 Video], 146 (original pagination). Section D.7.1.2 describes these video coding tools. The SNR scalability is summarized as : “The aim of SNR scalability is primarily to provide a mechanism for transmission of a two layer service, these two layers providing the same picture resolution but different quality level.” In brief, section F.7.1.2.1 explains how SNR scalability functions in the following way (emphasis added):

In the lower layer, the encoding is similar to the non scalable situation in terms of decisions, adaptive quantization, buffer regulation. The intra or error prediction macroblocks are DCT transformed. The coefficients are then quantized using a first rather coarse quantizer. The quantized coefficients are then VLC coded and sent together with the required

side information (macroblock_type, motion vectors, coded_block_pattern()).

In parallel, the quantized DCT coefficients coming from the lower layer, are de-quantized. The residual error between the coefficients and the de-quantized coefficients is then re-quantized, using a second finer quantizer. The resulting refinement coefficients are VLC coded and form the additional enhancement layer, together with a marginal amount of side information (quantiser_scale_code, coded_block_pattern() ...).

Ex.1035 [MPEG2/H.262 Video], 146 (original pagination).

102. The spatial scalability in MPEG2 is implemented by a variant of the SNR scalability in which the video may be decoded as a 2:1 reduced spatial (pixel) resolution.

103. Temporal scalability in MPEG2 is illustrated in the figure below and summarized in section D.7.4 as follows:

A two layer temporally scalable coding structure consisting of a base and an enhancement layer is shown in Figure F.1. Consider video input at full temporal rate to temporal demultiplexer; in our example it is temporally demultiplexed to form two video sequences, one input to the base layer encoder and the other input to the enhancement layer encoder. The base layer encoder is a non-hierarchical encoder operating at half temporal rate, the enhancement layer encoder is like a MAIN profile encoder and also operates at half temporal rate except that it uses base layer decoded pictures for motion compensated prediction. The encoded bitstreams of base and enhancement layers are multiplexed as a single stream in the systems multiplexer. The systems demultiplexer extracts two bitstreams and inputs corresponding bitstreams to base and enhancement layer decoders. The output of the base layer decoder can be shown standalone at half temporal rate or after multiplexing with enhancement layer decoded frames and shown at full temporal rate.

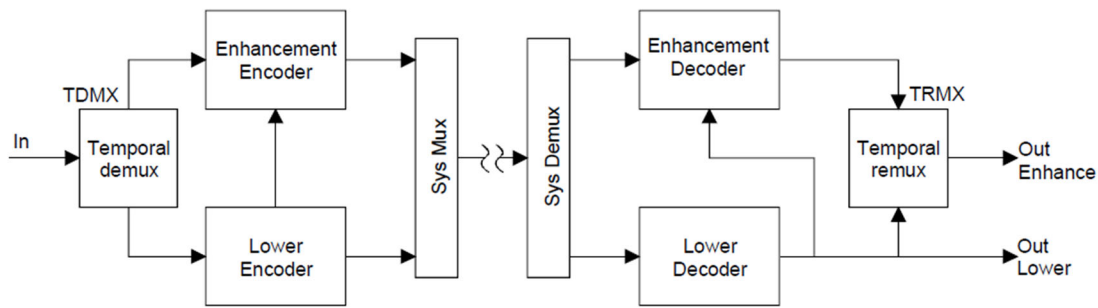


Figure D.1 – A two-layer codec structure for temporal scalability

Ex.1035 [MPEG2/H.262 Video], 147 (original pagination).

104. The MPEG4 Part 2 Visual standard includes fine grain scalability “FGS.” A comprehensive overview can be found in Radha et al, Ex.1040 (Radha, *The MPEG-4 Fine-Grained Scalable Video Coding Method*), which is incorporated by reference in this declaration in full. In brief, this comprises the following:

In order to meet the requirements outlined in the previous section, FGS encoding is designed to cover any desired bandwidth range while maintaining a very simple scalability structure. ... the FGS structure consists of only two layers: a base-layer coded at a bitrate and a single enhancement-layer coded using a fine-granular (or embedded) scheme to a maximum bitrate of R_e . This structure provides a very efficient, yet simple, level of abstraction between the encoding and streaming processes. The encoder only needs to know the range of bandwidth $[R_{min} - R_b, R_{max} = R_e]$ over which it has to code the content, and it does not need to be aware of the particular bitrate the content will be streamed at. The streaming server on the other hand has a total flexibility in sending any desired portion of any enhancement layer frame (in parallel with the corresponding base layer picture), without the need for performing complicated real-time rate control algorithms. This enables the server to handle a very large number of unicast streaming sessions and to adapt to their bandwidth variations in real-time. On the receiver side, the FGS framework adds a small amount of complexity and memory requirements to any standard motion-compensation based video decoder.

...
... the FGS framework requires two encoders, one for the base-layer and the other for the enhancement layer. The base-layer can be compressed using any motion-compensation video encoding method. Naturally, the DCT-based MPEG-4 video standard is a good candidate for the base-layer encoder due to its coding efficiency especially at low bitrates.

Using the DCT transform at both the base and enhancement layers enables the encoder to perform a simple residual computation of the FGS enhancement-layer. Each DCT FGS-residual frame consists of N_{BP} bitplanes:

$$N_{BP} = \lfloor \log_2 (|C|_{\max}) \rfloor + 1$$

Where $|C|_{\max}$ is the maximum DCT (magnitude) value of the residual frame under consideration.³ After identifying $|C|_{\max}$ and the corresponding N_{BP} , the FGS enhancement-layer encoder scans the residual signal using the traditional zig-zag scanning method starting from the most significant bitplane BP(1) and ending at the least significant bitplane BP(N_{BP}). Every bitplane consists of nonoverlapping macroblocks (MB's), and each MB includes four luminance () blocks and two chroma blocks (and). Run-length codes are used for (lossless) entropy-coding of the zeros and ones in each bitplane block. This process generates variable length codes that constitute the FGS compressed bitstream. A special "all-zero blocks" code is used when all six bitplane-blocks (within a given bitplane-macroblock) do not have any bits with a value of one. At the receiver side, the FGS bitstream is first decoded by a variable length decoder (VLD). Due to the embedded nature of the FGS stream, the VLD regenerates the DCT residual bitplanes starting from the most significant bitplane toward the least significant one. Moreover, due to the type of scanning used by the FGS encoder, it is possible that the decoder does not receive all of the bitplane-blocks that belong to a particular bitplane. Any bitplane block not received by the decoder can be filled with zero values. The resulting DCT residual is then inverse-transformed to generate the SNR residual pixels. These residual pixels are then added to the base-layer decoder output to generate the final enhanced scalable video.

Radha et al., Ex.1040 (Radha, *The MPEG-4 Fine-Grained Scalable Video Coding Method*).

6. The H.263 Standard

105. H.263 is a standard developed by ITU-T for videoconferencing applications. The work commenced in 1993, leading to three versions of the standard in 1996, 1998 and 2001. The first version is an evolution of the earlier H.261 standard. The second version, which is incorporated by reference in this declaration in full, included a number of optional annexes that are relevant to the current matter. See, e.g., Ex.1036 (ITU-T Recommendation H.263).

106. For example, Ex.1036, Annex J, Deblocking Filter Mode, specifies the optional use of a loop filter. The Introduction states: “This annex describes the use of an optional block edge filter within the coding loop. The main purpose of the block edge filter is to reduce blocking artifacts. ... The use of this mode is indicated in the PLUSPTYPE field of the picture header.” ITU-T Rec. H.263 (01/2005) p. 80.

107. In another example, Ex.1036, Annex O specifies three scalability modes, similar to MPEG2: SNR, spatial and temporal. In brief, section O.1.1 discloses temporal scalability as follows:

Temporal scalability is achieved using bidirectionally predicted pictures, or B-pictures. B-pictures allow prediction from either or both a previous and subsequent reconstructed picture in the reference layer. This property generally results in improved compression efficiency as compared to that of P-pictures. B-pictures (and the B-part of PB- or

Improved PB-frames) are not used as reference pictures for the prediction of any other pictures. This property allows for B-pictures to be discarded if necessary without adversely affecting any subsequent pictures, thus providing temporal scalability. Figure O.1 illustrates the predictive structure of P- and B-pictures.

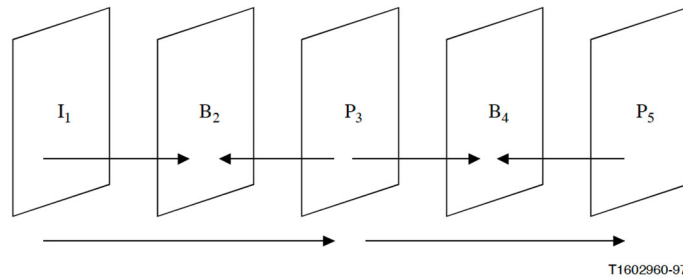


Figure O.1/H.263 – Illustration of B-picture prediction dependencies

Ex.1036 (ITU-T Recommendation H.263), 102-103 (original pagination).

108. SNR scalability is disclosed in section O.1.2 as follows:

The other basic method to achieve scalability is through spatial/SNR enhancement. Spatial scalability and SNR scalability are equivalent except for the use of interpolation as is described shortly. Because compression introduces artifacts and distortions, the difference between a reconstructed picture and its original in the encoder is (nearly always) a nonzero-valued picture, containing what can be called the coding error. Normally, this coding error is lost at the encoder and never recovered. With SNR scalability, these coding error pictures can also be encoded and sent to the decoder, producing an enhancement to the decoded picture. The extra data serves to increase the signal-to-noise ratio of the video picture, and hence, the term SNR scalability. Figure O.2 illustrates the data flow for SNR scalability. The vertical arrows from the lower layer illustrate that the picture in the enhancement layer is predicted from a reconstructed approximation of that picture in the reference (lower) layer. If prediction is only formed from the lower layer, then the enhancement layer picture is referred to as an EI-picture. It is possible, however, to create a modified bidirectionally predicted picture using both a prior enhancement layer picture and a temporally simultaneous lower layer reference picture. This type of picture is referred to as an EP-picture or "Enhancement" P-picture. The

prediction flow for EI- and EP-pictures is shown in Figure O.2. (Although not specifically shown in Figure O.2, an EI-picture in an enhancement layer may have a P-picture as its lower layer reference picture, and an EP-picture may have an I-picture as its lower-layer enhancement picture.) For both EI- and EP-pictures, the prediction from the reference layer uses no motion vectors. However, as with normal P-pictures, EP-pictures use motion vectors when predicting from their temporally prior reference picture in the same layer.

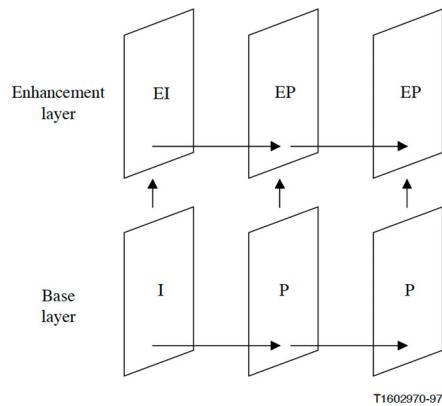


Figure O.2/H.263 – Illustration of SNR scalability

Ex.1036 (ITU-T Recommendation H.263), 103-104 (original pagination).

109. Spatial scalability is disclosed in section O.1.3 as follows:

The third and final scalability method in the Temporal, SNR, and Spatial Scalability mode is spatial scalability, which is closely related to SNR scalability. The only difference is that before the picture in the reference layer is used to predict the picture in the spatial enhancement layer, it is interpolated by a factor of two either horizontally or vertically (1-D spatial scalability), or both horizontally and vertically (2-D spatial scalability).

Other than requiring an upsampling process to increase the size of the reference layer picture prior to its use as a reference for the encoding process, the processing and syntax for a spatial scalability picture is functionally identical to that for an SNR scalability picture. Since there is very little syntactical distinction between pictures using SNR scalability and pictures using spatial scalability, the pictures used for either purpose are called EI- and EP-pictures. The picture in the base

layer which is used for upward prediction in an EI- or EP-picture may be an I-picture, a P-picture, or the P-part of a PB- or Improved PB-frame (but shall not be a B-picture or the B-part of a PB- or Improved PB-frame).

Ex.1036 (ITU-T Recommendation H.263), 104-105 (original pagination).

110. Spatial scalability is also provided by two other Annexes. Annex P specifies a Reference Picture Resampling, while Annex Q specifies a Reduced Resolution Update. Annex P Section P.1 Introduction of Ex.1036 explains:

This annex describes the use and syntax of a resampling process which can be applied to the previous decoded reference picture in order to generate a "warped" picture for use in predicting the current picture. This resampling syntax can specify the relationship of the current picture to a prior picture having a different source format, and can also specify a "global motion" warping alteration of the shape, size, and location of the prior picture with respect to the current picture. In particular, the Reference Picture Resampling mode can be used to adaptively alter the resolution of pictures during encoding.

Ex.1036 (ITU-T Recommendation H.263), 116 (original pagination).

111. Section P.5 Factor-of-4 Resampling further specifies:

Factor-of-4 resampling, which converts both the horizontal and vertical size of the picture by a factor of 2 or $\frac{1}{2}$, is a special case of the resampling algorithm described in P.3. The simplified description of the resampling algorithm for this special case is provided in this subclause.

Ex.1036 (ITU-T Recommendation H.263), 126 (original pagination).

112. Annex Q Section Q.1 Introduction explains:

This annex describes an optional Reduced-Resolution Update mode of this Recommendation. The capability of this mode is signaled by

external means (for example, Recommendation H.245). The use of this mode is indicated in the PLUSPTYPE field of the picture header. The Reduced-Resolution Update mode is expected to be used when encoding a highly active scene, and provides the opportunity to increase the coding picture rate while maintaining sufficient subjective quality. This mode allows the encoder to send update information for a picture that is encoded at a reduced resolution, while preserving the detail in a higher resolution reference image to create a final image at the higher resolution.

Ex.1036 (ITU-T Recommendation H.263), 129 (original pagination).

113. All of these Annexes can be used in conjunction with Annex N, which specifies the use of multiple reference pictures and the capability to recover from a transmission error by transmitting a new reference picture (or portion of a new reference picture) when the decoder receives a corrupted picture. Section N.2 provides an introduction:

The source coder of this mode is shown in generalized form in Figure N.1. This figure shows a structure which uses a number of picture memories. The source coder may select one of the picture memories to suppress the temporal error propagation due to the inter-frame coding. The information to signal which picture is selected for prediction is included in the encoded bitstream.

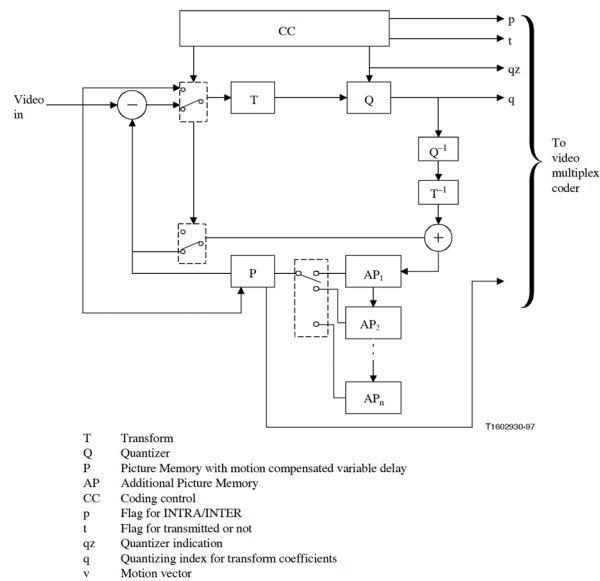


Figure N.1/H.263 – Source coder for NEWPRED

Ex.1036 (ITU-T Recommendation H.263), 97 (original pagination).

114. Upon detecting an error in a received picture, the decoder may use a backchannel to signal the encoder to send a new picture. The encoder may choose to send an I-picture, so the decoder can reinitialize its temporal prediction of pictures and recover from the error propagation. Because I-pictures require a significant number of transmitted bits, the encoder may combine Annex N with any of the mechanisms in Annexes O, P, and Q to transmit the I-picture at 2:1 reduced resolution. (Transmitting a full-resolution I-picture may require 5-10 times as many bits as a P-picture meaning it could take 5-10 picture display periods. The encoder and decoder rate buffers are sized to manage the variable rate at which pictures are transmitted.)

115. The new reference picture is signaled using its temporal reference. In the normal mode of operation, the most recent temporally previous picture is used as the reference picture. Section N.4.1 specifies a flag Temporal Reference for Prediction Indicator (TRPI), which when set causes a Temporal Reference for Prediction (TRP) to be transmitted. Section N.4.1.4 specifies:

When present (as indicated in TRPI), TRP indicates the Temporal Reference which is used for prediction of the encoding, except in the case of B-pictures and the B-picture part of an Improved PB-frame. For B-pictures or the B-picture part of an Improved PB-frame, the picture having the temporal reference TRP is used for the prediction in the forward direction. (Prediction in the reverse temporal direction always uses the immediately temporally subsequent picture.) TRP is a ten-bit number. If a custom picture clock frequency was not in use for the reference picture, the two MSBs of TRP are zero and the LSBs contain the eight-bit TR found in the picture header of the reference picture. If a custom picture clock frequency was in use for the reference picture, TRP is a ten-bit number consisting of the concatenation of ETR and TR from the reference picture header. When TRP is not present, the most recent temporally previous anchor picture shall be used for prediction, as when not in the Reference Picture Selection mode. TRP is valid until the next PSC, GSC, or SSC.

Ex.1036 (ITU-T Recommendation H.263), 99 (original pagination).

116. The third version of H.263 includes Annex U – Enhanced Reference Picture Selection mode – published initially in 2000 as a separate ITU document¹¹. ERPS mode provides more flexible motion compensated prediction with the

¹¹ ITU-T H.263/Annex U (11/2000).

freedom to choose the prediction of a current picture, or block within a current picture from multiple reference pictures. *See e.g.*, U.1 Introduction, p.1 (emphasis added):

For coding efficiency, motion compensation can be extended to prediction from multiple pictures. The extension of motion compensation to multi-picture prediction is achieved by **extending each motion vector by a picture reference parameter** that is used to address a macroblock or block prediction region for motion compensation in any of the multiple reference pictures. **The picture reference parameter is a variable length code specifying a relative buffer index.** The reference pictures are assembled in a buffering scheme that is controlled by the encoder.

117. The standard provides mechanisms by which as successive pictures are encoded, transmitted, received and decoded, a decision is made whether the encoder and decoder should store the decoded picture in their respective picture buffers for future use as a reference picture for coding a future picture. A set of picture memories is maintained for this purpose (see Figure U.1 below), and Annex U ERPS Mode includes mechanisms for reference frame indexing and reference buffer management.

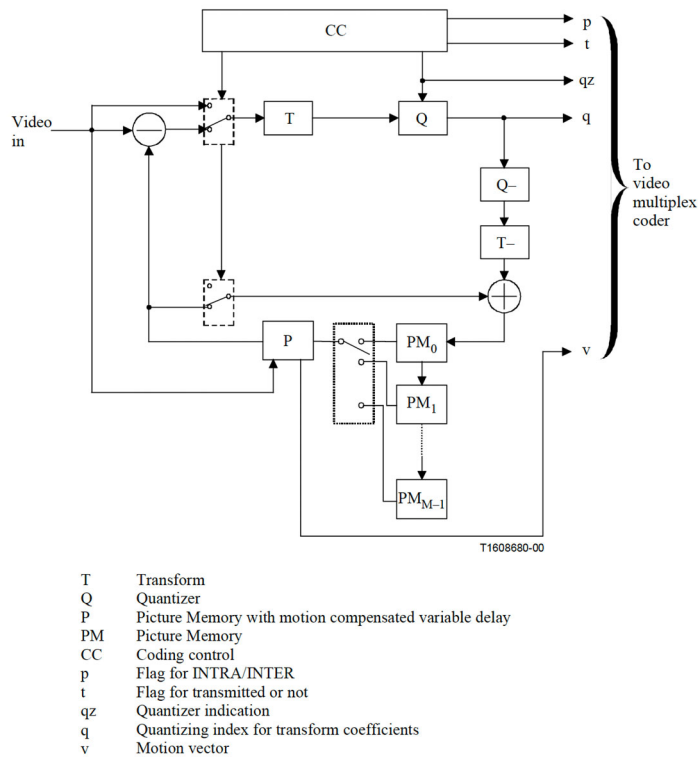


Figure U.1/H.263 – Source coder for Enhanced Reference Picture Selection mode

118. Signaling of the optional H.263 video coding tools (annexes) is specified using a syntax construct called PLUSPTYPE, see H.263 clause 5.1. The use of optional tools is bit-encoded in a word of the PLUSRTYPE syntax called OPPTYPE – see clause 5.1.4.2, which is reproduced below:

5.1.4.2 The Optional Part of PLUSPTYPE (OPPTYPE) (18 bits)

If UFEP is "001", then the following bits are present in PLUSPTYPE:

- Bits 1-3 Source Format, "000" reserved, "001" sub-QCIF, "010" QCIF, "011" CIF, "100" 4CIF, "101" 16CIF, "110" custom source format, "111" reserved;
- Bit 4 Optional Custom PCF, "0" CIF PCF, "1" custom PCF;
- Bit 5 Optional Unrestricted Motion Vector (UMV) mode (see Annex D), "0" off, "1" on;
- Bit 6 Optional Syntax-based Arithmetic Coding (SAC) mode (see Annex E), "0" off, "1" on;
- Bit 7 Optional Advanced Prediction (AP) mode (see Annex F), "0" off, "1" on;
- Bit 8 Optional Advanced INTRA Coding (AIC) mode (see Annex I), "0" off, "1" on;
- Bit 9 Optional Deblocking Filter (DF) mode (see Annex J), "0" off, "1" on;
- Bit 10 Optional Slice Structured (SS) mode (see Annex K), "0" off, "1" on;
- Bit 11 Optional Reference Picture Selection (RPS) mode (see Annex N), "0" off, "1" on;
- Bit 12 Optional Independent Segment Decoding (ISD) mode (see Annex R), "0" off, "1" on;
- Bit 13 Optional Alternative INTER VLC (AIV) mode (see Annex S), "0" off, "1" on;
- Bit 14 Optional Modified Quantization (MQ) mode (see Annex T), "0" off, "1" on;
- Bit 15 Equal to "1" to prevent start code emulation;
- Bit 16 Reserved, shall be equal to "0";
- Bit 17 Reserved, shall be equal to "0";
- Bit 18 Reserved, shall be equal to "0".

119. It can be seen for example the loop filter (Annex J) can be turned ON and OFF by setting Bit 9. Similar syntax provides for turning the optional Reduced Resolution Update mode (Annex Q) ON and OFF.

7. The H.264 Standard

120. H.264 is a standard jointly developed by ITU-T for videoconferencing applications and ISO for consumer video applications. (The ISO designation is 14496-10, i.e., MPEG4 Part 10 also known as AVC – Advanced Video Codec, Ex.1032). The work commenced in 1998 in the ITU-T workgroup VCEG (Video

Coding Experts Group), with MPEG joining in December 2001 to form the Joint Video Team – JVT. The first version was published in May 2003.

121. H.264/AVC provides a significant improvement in video coding performance over MPEG2 and is still the most widely used video coding standard worldwide. Notably it is used for High Definition Television broadcasting (HDTV). The multiple reference picture capabilities of H.263 Annex U were incorporated into H.264 (except a sub-picture mode). In H.264/AVC a coded block may include coded motion data that comprises a motion vector and a reference index that selects which one of multiple reference frames is used.

122. Scalable Video Coding (SVC) was included from version 3 in November 2007 (ITU-T Rec. H.264 (11/2007), Ex.1032 (*ITU-T Recommendation H.264*)), which is incorporated by reference in this declaration in full.

123. For example, the Schwarz IEEE paper “Overview of the scalable video coding extension of the H. 264 AVC standard” by three of the principle contributors to the H.264 scalability specification, Ex.1010 (Schwarz, *Overview of the Scalable Video Coding Extension*), which is incorporated by reference in this declaration in full, provides a thorough overview of the practical applications facilitated by the H.264 improvements.

124. Like the earlier standards, H.264/AVC supports the three types of scalability – temporal, spatial and quality (SNR).

125. Temporal scalability may be implemented using hierarchical B-Pictures. In H.264/AVC B-pictures can be used as reference pictures. Figure 1 of the paper illustrates different modes:

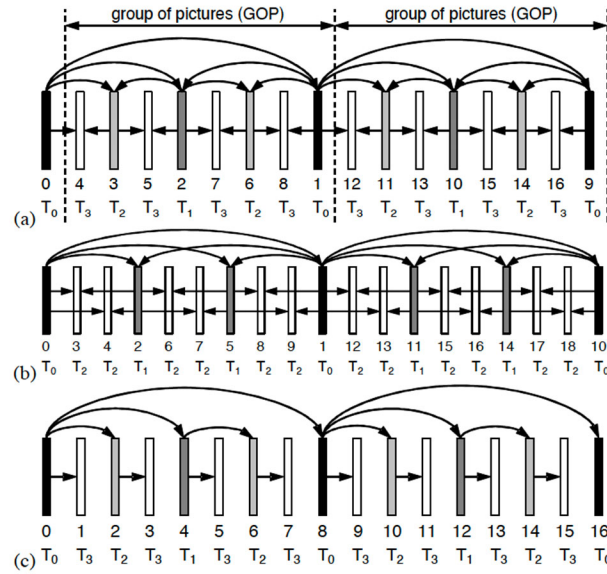


Fig. 1. Hierarchical prediction structures for enabling temporal scalability: (a) coding with hierarchical B pictures, (b) non-dyadic hierarchical prediction structure, (c) hierarchical prediction structure with a structural encoder/decoder delay of zero. The numbers directly below the pictures specify the coding order, the symbols T_k specify the temporal layers with k representing the corresponding temporal layer identifier.

Ex.1010 (Schwarz, *Overview of the Scalable Video Coding Extension*), FIG. 1.

126. Note the pictures are arranged as GOPs in which base-layer pictures demarcate the GOP boundaries.

Temporal scalability with dyadic temporal enhancement layers can be very efficiently provided with the concept of hierarchical B pictures [28][29] as illustrated in Fig. 1a2. The enhancement layer pictures are typically coded as B pictures, where the reference picture lists 0 and 1 are restricted to the temporally preceding and succeeding picture, respectively, with a temporal layer identifier less than the temporal layer identifier of the predicted picture. Each set of temporal layers $\{T_0, \dots, T_k\}$ can be decoded independently of all layers with a temporal

layer identifier $T > k$. In the following, the set of pictures between two successive pictures of the temporal base layer together with the succeeding base layer picture is referred to as a group of pictures (GOP). Although the described prediction structure with hierarchical B pictures provides temporal scalability and also shows excellent coding efficiency as will be demonstrated later, it represents a special case. In general, hierarchical prediction structures for enabling temporal scalability can always be combined with the multiple reference picture concept of H.264/AVC. This means that the reference picture lists can be constructed by using more than one reference picture, and they can also include pictures with the same temporal level as the picture to be predicted. Furthermore, hierarchical prediction structures are not restricted to the dyadic case. As an example, Fig. 1b illustrates a non-dyadic hierarchical prediction structure, which provides 2 independently decodable sub-sequences with 1/9-th and 1/3-rd of the full frame rate. It should further be noted that it is possible to arbitrarily modify the prediction structure of the temporal base layer, e.g., in order to increase the coding efficiency. The chosen temporal prediction structure does not need to be constant over time. Note that it is possible to arbitrarily adjust the structural delay between encoding and decoding a picture by restricting motion-compensated prediction from pictures that follow the picture to be predicted in display order. As an example, Fig. 1c shows a hierarchical prediction structure, which does not employ motion-compensated prediction from pictures in the future. Although this structure provides the same degree of temporal scalability as the prediction structure of Fig. 1a, its structural delay is equal to zero compared to 7 pictures for the prediction structure in Fig. 1a. However, such low-delay structures typically decrease coding efficiency.

Ex.1010 (Schwarz, *Overview of the Scalable Video Coding Extension*).

127. Spatial scalability in H.264/AVC is an enhanced version of the spatial scalability in the earlier standards. It is illustrated in Figure 4, which includes a combination with temporal scalability:

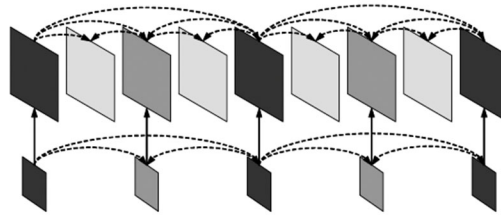


Fig. 4. Multi-layer structure with additional inter-layer prediction for enabling spatial scalable coding.

For supporting spatial scalable coding, SVC follows the conventional approach of multi-layer coding, which is also used in H.262 | MPEG-2 Video, H.263, and MPEG-4 Visual. Each layer corresponds to a supported spatial resolution and is referred to by a spatial layer or dependency identifier D . The dependency identifier D for the base layer is equal to 0, and it is increased by 1 from one spatial layer to the next. In each spatial layer, motion-compensated prediction and intra prediction are employed as for single-layer coding. But in order to improve coding efficiency in comparison to simulcasting different spatial resolutions, additional so-called inter-layer prediction mechanisms are incorporated as illustrated in Fig. 4.

In order to restrict the memory requirements and decoder complexity, SVC specifies that the same coding order is used for all supported spatial layers. The representations with different spatial resolutions for a given time instant form an access unit and have to be transmitted successively in increasing order of their corresponding spatial layer identifiers D . But as illustrated in Fig. 4, lower layer pictures do not need to be present in all access units, which makes it possible to combine temporal and spatial scalability.

Ex.1010 (Schwarz, *Overview of the Scalable Video Coding Extension*).

128. Quality scalability in the H.264/AVC standard is provided in several modes, one of which is called coarse grain scalability – CGS. The paper explains:

Quality scalability can be considered as a special case of spatial scalability with identical picture sizes for base and enhancement layer. As already mentioned in sec. V.B, this case is supported by the general

concept for spatial scalable coding and it is also referred to as coarse-grain quality scalable coding (CGS). The same inter-layer prediction mechanisms as for spatial scalable coding are employed, but without using the corresponding upsampling operations and the inter-layer deblocking for intra-coded reference layer macroblocks. Furthermore, the inter-layer intra and residual prediction are directly performed in the transform domain. When utilizing inter-layer prediction for coarse-grain quality scalability in SVC, a refinement of texture information is typically achieved by requantizing the residual texture signal in the enhancement layer with a smaller quantization step size relative to that used for the preceding CGS layer.

Ex.1010 (Schwarz, *Overview of the Scalable Video Coding Extension*).

129. The quality scalable modes deployed by the legacy standards and H.264/AVC are illustrated in Figure 8:

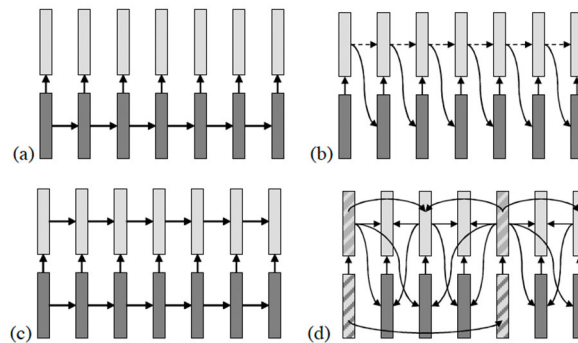


Fig. 8. Various concepts for trading off enhancement layer coding efficiency and drift for packet-based quality scalable coding: (a) base layer only control, (b) enhancement layer only control, (c) two-loop control, (d) key picture concept of SVC for hierarchical prediction structures, where key pictures are marked by the hatched boxes.

Ex.1010 (Schwarz, *Overview of the Scalable Video Coding Extension*).

130. Schwarz explains the differences between these modes in detail, but in brief, the tradeoff issues include drift in the feedback loop of the decoder, consideration of transmission error conditions, and the cost of providing error resilience in terms of coding efficiency.

131. H.264 includes an optional tool for switching efficiently between contemporaneous coded streams. For example, two streams may comprise the same content with sequences of the pictures coded at different bitrates, and the functionality is desired to switch from one bitrate to the other bitrate seamlessly. While that could be accomplished with periodic I-pictures, such pictures would require a significant number of coded bits and attendant loss of coding efficiency. H.264 provides a tool at the Slice level called SP Slices, in which the slices are predictively coded with fewer bits in such a way that either stream may continue error-free decoding from the SP Slice. Of course, such a Slice may cover the entire Picture if desired. This tool is disclosed in a white paper - H.264 / MPEG-4 Part 10: SP and SI slices (Ex.1041 (*H.264/MPEG-4 Part 10 White Paper*)). In brief:

2. SP and SI slices

SP and SI slices are specially-coded slices that enable (among other things) efficient switching between video streams and efficient random access for video decoders. A common requirement is for a video decoder to switch between one of several encoded streams. For example, the same video material is coded at multiple bitrates for transmission across the Internet; a decoder attempts to decode the highest-bitrate stream it can receive but may require to switch automatically to a lower-bitrate stream if the data throughput drops.

Example: A decoder is decoding Stream A and wants to switch to decoding Stream B (Figure 2-1). For simplicity, assume that each frame is encoded as a single slice and predicted from one reference (the previous decoded frame). After decoding P-slices A0 and A1, the decoder wants to switch to Stream B and decode B2, B3 and so on. If all the slices in Stream B are coded as P-slices, then the decoder will not have the correct decoded reference frame(s) required to reconstruct

B2 (since B2 is predicted from the decoded frame B1 which does not exist in stream A). One solution is to code frame B2 as an I-slice. Because it is coded without prediction from any other frame, it can be decoded independently of preceding frames in stream B. The decoder can therefore switch between stream A and stream B as shown in Figure 2-1. Switching can be accomplished by inserting an I-slice at regular intervals in the coded sequence to create “switching points”. However, an I-slice is likely to contain much more coded data than a P-slice and the result is a peak in the coded bitrate at each switching point.

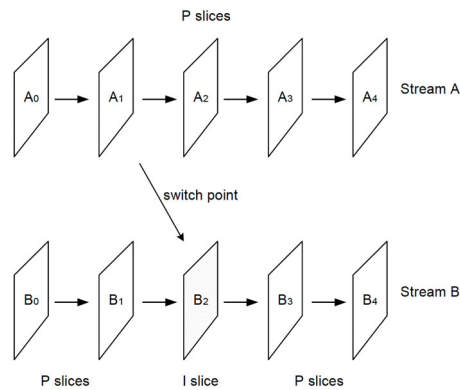


Figure 2-1 Switching streams using I-slices

SP-slices are designed to support switching without the increased bitrate penalty of I-slices. Figure 2-2 shows how they may be used. At the switching point (frame 2 in each sequence), there are now 3 coded SP-slices. These are coded using motion compensated prediction (making them more efficient than I-slices). SP-slice A2 can be decoded using reference frame A1; SP-slice B2 can be decoded using reference frame B2. The key to the switching process is SP-slice AB2 (known as a switching SP-slice): this is created in such a way that it can be decoded using motion-compensated reference frame A1, to produce decoded frame B2 (i.e. the decoder output frame B2 is identical whether decoding B1 followed by B2 or A1 followed by AB2). An extra SP-slice is required at each switching point (and in fact another SP-slice, BA2, would be required to switch in the other direction) but this is still more efficient than encoding frames A2 and B2 as I-slices.

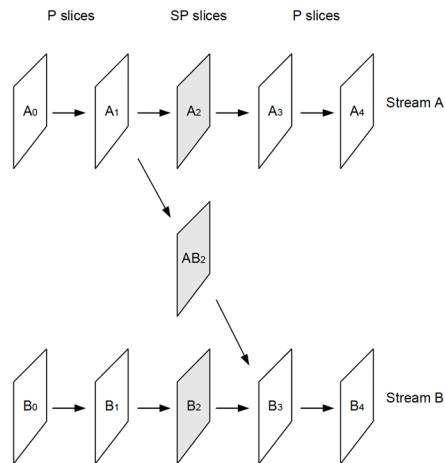


Figure 2-2 Switching streams using SP-slices

Ex.1041 (*H.264/MPEG-4 Part 10 White Paper*).

132. Of course one way to encode Stream A at one bitrate and Stream B at another bitrate is to use a different resolution for each stream.

133. A comprehensive overview can be found in Karczewicz M, et al. Ex.1033, which is incorporated by reference in this declaration in full.

134. H.264 specifies a mechanism called the Network Abstraction Layer – NAL – for encapsulating the coded video data and associated parametric data into data units that may be stored or transmitted.

3.89 NAL unit: A syntax structure containing an indication of the type of data to follow and bytes containing that data in the form of an RBSP interspersed as necessary with emulation prevention bytes.

Ex.1032 (*ITU-T Recommendation H.264*) clause 3.97

Scalable video coding is specified in Annex G allowing the construction of bitstreams that contain sub-bitstreams that conform to this Specification. For temporal bitstream scalability, i.e., the presence of a sub-bitstream with a smaller temporal sampling rate than the bitstream, complete access units are removed from the bitstream when

deriving the sub-bitstream. In this case, high-level syntax and inter prediction reference pictures in the bitstream are constructed accordingly. For spatial and quality bitstream scalability, i.e., the presence of a sub-bitstream with lower spatial resolution or quality than the bitstream, NAL units are removed from the bitstream when deriving the sub-bitstream. In this case, inter-layer prediction, i.e., the prediction of the higher spatial resolution or quality signal by data of the lower spatial resolution or quality signal, is typically used for efficient coding.

Ex.1032 (*ITU-T Recommendation H.264*).

III. Relevant Legal Standards

135. I have been asked to provide my opinions as to whether claims 1-3 of the '448 Patent would have been obvious to a person of ordinary skill in the art as of the earliest claimed priority date of the '448 Patent (May 30, 2008) ("Critical Date").

136. I am an engineer by training and profession. The opinions I express in this declaration involve the application of my technical knowledge and experience to the evaluation of certain prior art with respect to the '448 Patent. In addition, I understand that the following legal principles apply.

137. It is my understanding that, in determining whether claims of the '448 Patent are obvious in this proceeding, the claim terms are generally given their ordinary and customary meaning as understood by a person of ordinary skill in the relevant art. A person of ordinary skill in the art would read the claim terms in the context of the entire patent specification in which they appear, as well as the prosecution history of the patent.

138. It is my understanding that a claim is unpatentable under 35 U.S.C. § 103 if the claimed subject matter as a whole would have been obvious to a person of ordinary skill in the art at the time of the alleged invention. I also understand that an obviousness analysis takes into account the scope and content of the prior art, the differences between the claimed subject matter and the prior art, and the level of ordinary skill in the art at the time of the invention.

139. In determining the scope and content of the prior art, it is my understanding that a reference is considered relevant prior art if it falls within the field of the inventor's endeavor. In addition, a reference is prior art if it is reasonably pertinent to the particular problem with which the inventor was involved. A reference is reasonably pertinent if it logically would have commended itself to an inventor's attention in considering his problem. If a reference relates to the same problem as the claimed invention, that supports use of the reference as prior art in an obviousness analysis.

140. To assess the differences between prior art and the claimed subject matter, it is my understanding that 35 U.S.C. § 103 requires the claimed invention to be considered as a whole. This "as a whole" assessment involves showing that one of ordinary skill in the art at the time of invention, confronted by the same problems as the inventor and with no knowledge of the claimed invention, would

have selected the elements from the prior art and combined them in the claimed manner.

141. It is my further understanding that several rationales may be applied for combining references or modifying a reference to show obviousness of claimed subject matter. These rationales include: combining prior art elements according to known methods to yield predictable results; simple substitution of one known element for another to obtain predictable results; a predictable use of prior art elements according to their established functions; applying a known technique to a known device (method or product) ready for improvement to yield predictable results; choosing from a finite number of identified, predictable solutions, with a reasonable expectation of success; and some teaching, suggestion, or motivation in the prior art that would have led one of ordinary skill to modify a prior art reference or to combine prior art teachings to arrive at the claimed invention.

IV. Person of Ordinary Skill in the Art

142. It is my understanding that when interpreting the claims of the '448 Patent I must do so based on the perspective of one of ordinary skill in the art at the time of the alleged invention ("POSITA"). For purposes of my analysis in this declaration, I have taken the earliest claimed priority date of the '448 Patent (May 30, 2008) as the date of the alleged invention ("Critical Date"). I understand that the factors considered in determining the ordinary level of skill in a field of art include

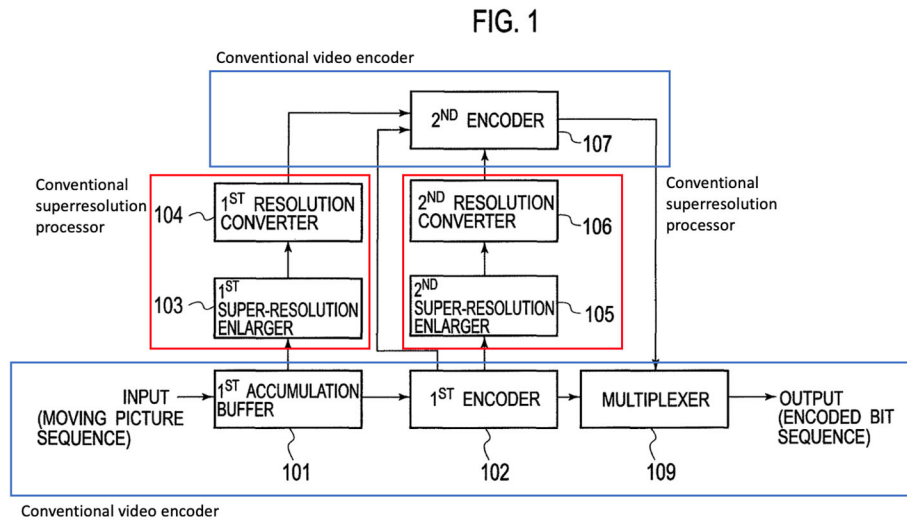
the level of education and experience of persons working in the field; the types of problems encountered in the field; the teachings of the prior art, and the sophistication of the technology at the time of the alleged invention. I understand that a POSITA is not a specific real individual, but rather is a hypothetical individual having the qualities reflected by the factors above. I understand that a POSITA would also have knowledge from the teachings of the prior art, including the art cited below.

143. Taking these factors into consideration, it is my opinion that one of ordinary skill in the art in the field of digital video coding as of the Critical Date, would have had a bachelor's degree in computer science, electrical engineering, computer engineering, or a related field, with 2-3 years of experience in digital video processing, or a related field. Additional graduate education could substitute for professional experience, or additional professional experience could substitute for formal education.

144. I have possessed the qualifications of a POSITA since the Critical Date of the '448 Patent, and long before.

V. Overview of the Subject Patent 9,042,448

145. The '448 Patent originated from a PCT application filed May 28, 2009, which is derived from a pair of Japanese applications, the earliest of which was filed May 30, 2008. Ex.1001 ('448 Patent), Cover.

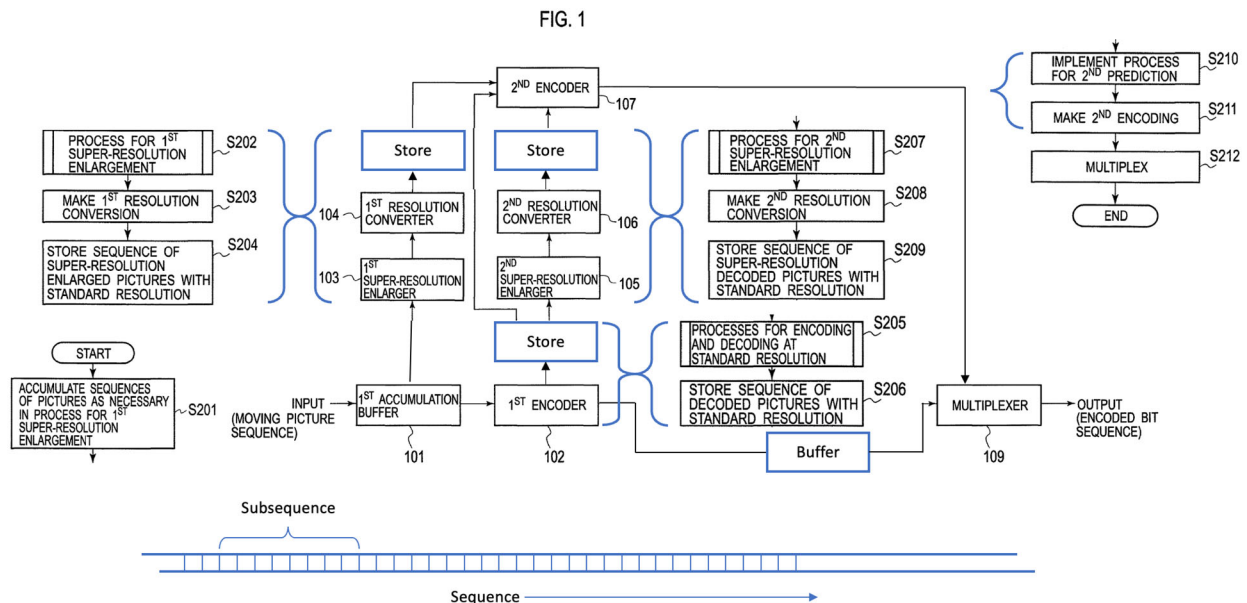


146. Figure 1 (annotated) illustrates an encoding system for video pictures. The 1st encoder is a conventional video encoder that inputs a succession of video pictures from the 1st accumulation buffer, encodes the pictures and outputs an encoded bit sequence (via a multiplexer.) The 1st encoder conventionally decodes each picture, stores the decoded pictures and provides them as “reference” pictures for coding of subsequent pictures. The 2nd encoder is a conventional video encoder, that inputs a succession of “target” video pictures from the 1st accumulation buffer (via a superresolution processor), encodes the target pictures and outputs an encoded bit sequence (via the multiplexer). The 2nd encoder employs temporal prediction conventionally selectively using one of two reference pictures derived from the output decoded pictures of the 1st encoder. The multiplexer selects the coded bitstream from either the 1st or 2nd encoder.

147. The 1st superresolution enlarger and 1st resolution converter collectively comprise a conventional superresolution processor that inputs a subsequence of pictures at a given pixel resolution and outputs a (single) picture at the same pixel resolution, but with enhanced information resolution (higher spatial frequency content). This output provides the target picture for the 2nd encoder. The 2nd superresolution enlarger and 2nd resolution converter collectively comprise a conventional superresolution processor that inputs a subsequence of pictures at a given pixel resolution that were decoded and output by the first decoder and outputs a (single) picture at the same pixel resolution, but with enhanced information resolution (higher spatial frequency content). This output provides one of the two reference picture for the 2nd encoder. The other reference picture is the unenhanced same picture decoded and output from the 1st encoder.

148. The figure illustrates a functional flow of processing, but does not illustrate either an instant in time for the processing of a given video picture, or a temporal flow for the processing of a succession of pictures. The figure does not show all the buffering that is required, first in support of the superresolution processing and second for encoding the target picture using the reference pictures. With reference to Figure 2 and the Technical Background section of this declaration, the temporal flow will be explained.

149. First, it should be clear that 1st Encoder and 2nd Encoder may not process the same input video picture at the same time instant, i.e., the same picture interval. Rather, the 1st encoder may produce reference pictures at an earlier picture interval for use by the 2nd encoder in a current picture interval. This is standard practice in video encoding, see for example H.263, 1998. Second, it should be clear the superresolution processors employ a subsequence of video pictures at the input to generate a single picture at the output. Figure 2 illustrates some of the “stores” needed to buffer video pictures and provide delays. The figure below combines figures 1 and 2 to provide a more complete picture of the encoding process¹².



150. The video pictures to be encoded form a sequence of pictures. The embodiment operates using subsequences of the sequence, but the specification and

¹² See also Figure 10.

claims of the '448 Patent are inconsistent in the use of terminology – the word “sequence” may refer to all the video pictures entering the 1st accumulation buffer 101 and the encoded bitstream representing those pictures exiting the multiplexer 109. But it may also refer to a subsequence of pictures from that sequence that are used during the processing of each picture of the sequence. The same lack of clarity is found in the patent’s use of the term “set of pictures”.

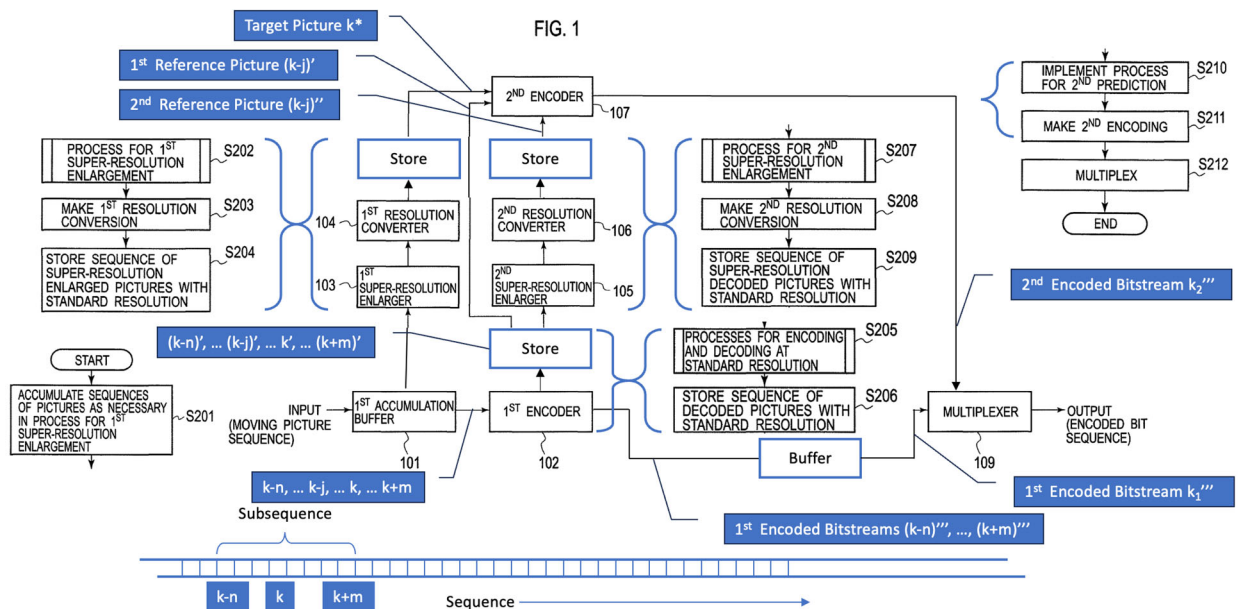
151. Let us consider the operations performed to process one picture from the sequence of pictures. Let us call this the target picture. (Such operations will be performed in due course for each of the pictures in the sequence of pictures). In my opinion, element S201 accumulates a *subsequence* of pictures to support the superresolution processing for a single, target picture. The 1st encoder encodes this subsequence of pictures, entropy encodes the pictures (via an “unshown entropy encoder”), passes the encoded bitstreams to a buffer prior to multiplexer 109 and stores the decoded pictures in the store S206. The accumulated subsequence of pictures are also superresolution processed by the 1st superresolution processor and 1st resolution converter, and a superresolution processed picture is stored in store S204. This picture will be the target picture to be encoded by the 2nd encoder. The stored subsequence of decoded pictures from the 1st encoder are also superresolution processed by the 2nd superresolution processor and 2nd resolution converter, and a superresolution processed picture is stored in store S204. This picture will be a first

reference picture to be used selectively by the 2nd encoder. The corresponding picture from the subsequence of decoded pictures in store 206 will be used selectively as a second reference picture. The second encoder encodes the target picture, selectively using one of the two reference pictures. The selectivity may be performed on what the specification calls “regions”, formed by dividing the target picture and the reference pictures into “no spaces left regions”. The encoded picture from the 2nd encoder is entropy encoded (via an “unshown entropy encoder”) into an encoded bitstream for the target picture and passes to the multiplexer 109. The 2nd encoder also passes information on which of the two reference pictures was selected to an “unshown entropy encoder” that passes the encoded data to the multiplexer 109. Parametric data used by the encoders may also be entropy encoded and passed to the multiplexer 109. The multiplexer multiplexes all of the passed information into a bitstream based on a syntax specification.

152. As an example, for encoding a current “target” frame k , the following temporal sequence of operations is required:

- 1st accumulation buffer 101 accumulates a subsequence of pictures, e.g., $k-n$, ... k , ... $k+m$
- 1st encoder encodes the subsequence of pictures, outputs the sequence of coded bitstreams $(k-n)''$, ..., $(k+m)''$ via an unshown entropy encoder to a buffer before the multiplexer 109 and outputs the decoded pictures, including a picture $(k-j)'$ that will be used as reference pictures $(k-n)'$, ... $(k-j)'$, ... k' , ... $(k+m)'$ to a store S206

- 1st superresolution processor and 1st resolution converter process the subsequence of pictures, and output the superresolution processed target picture k^* to a store S204
- 2nd superresolution processor and 2nd resolution converter process the subsequence of decoded pictures, and output a reference picture $(k-j)''$ to a store S209
- 2nd encoder encodes the superresolution processed target picture k^* selectively employing decoded picture $(k-j)'$ and superresolution processed decoded picture $(k-j)''$ as reference pictures
- The 2nd encoder outputs 2nd encoded bitstream k_2''' and entropy encoded selection information to the multiplexer 109 via an unshown entropy encoder
- The buffer after the 1st encoder outputs 1st encoded bitstream k_1''' to the multiplexer 109 via an unshown entropy encoder
- The multiplexer employs a syntax to multiplex the 1st and 2nd encoded bitstreams, the encoded selection information and also other parametric data



A. Overview

153. The '448 Patent specification and claims are written in language that can appear difficult to understand. Many phrases and sentences use terms that are not used in the art, and in particular certain words are overloaded.

154. For clarity, it must be understood that certain words have different apparent meanings in different limitations of the same claim and even within a limitation of a claim. A glossary provided here introduces the terms and the multiple such meanings.

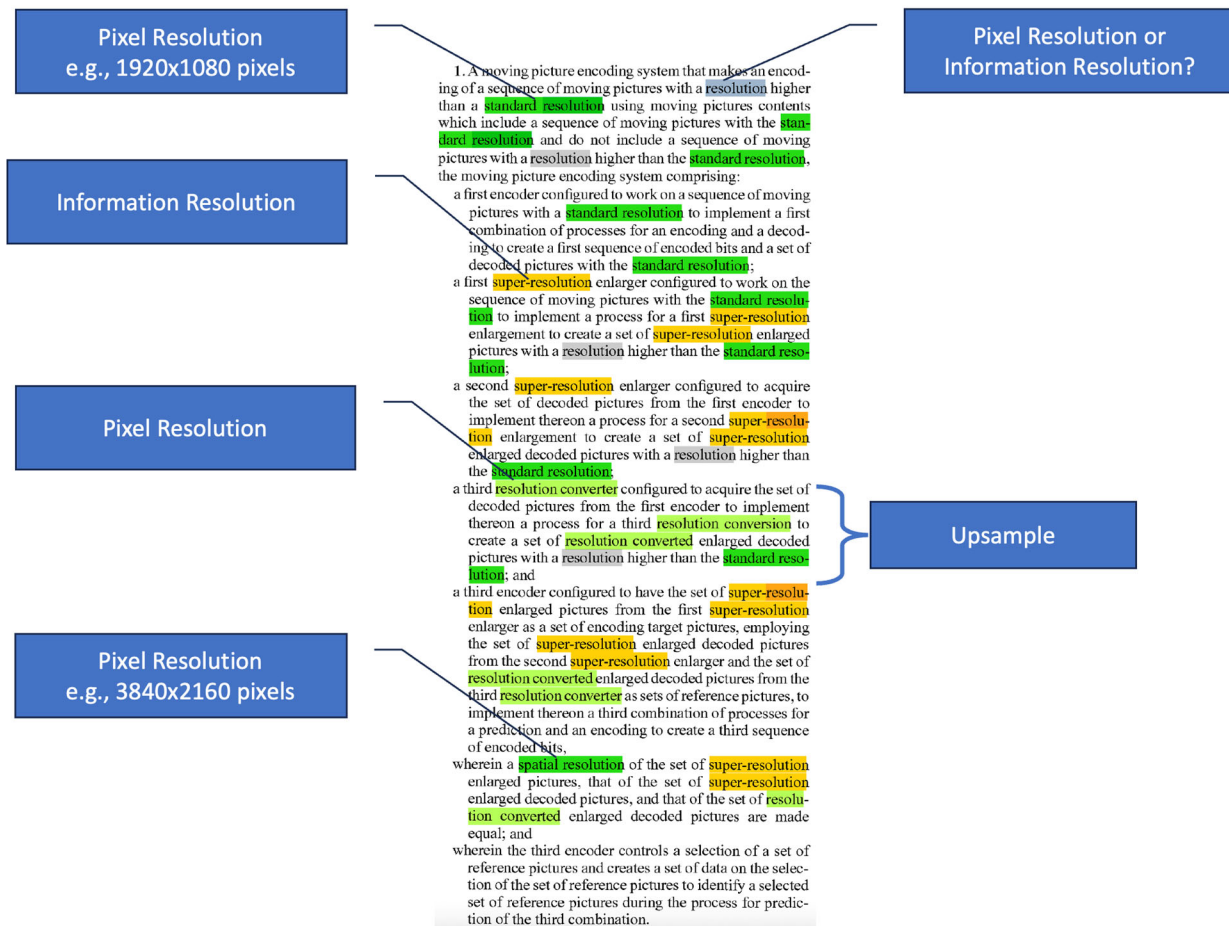
1. Resolution; Claims 1-3

155. This is an overloaded term in the art. The first usage of the term denotes the size of an image or video picture as an array of numbers of pixels. Historically, analog video “resolution” was expressed as the number of vertical “lines”, for example, US NTSC was 525-line video, while European PAL was 625-line video. More recently, US HDTV is broadcast primarily in two formats called “1080i” and “720p” again referencing the number of vertical “lines” which since the format is digital, means the number of pixels. The two formats comprise pixel arrays of 1920 x 1080 pixels and 1280 x 720 pixels respectively. Now there is an ultra-high definition format, having twice the number of pixels horizontally and vertically as the 1080i format, i.e., 3840 x 2160 pixels. The industry has decided to call this “4K”, i.e., using the number of pixels horizontally, rather than vertically. This usage could

be called “pixel resolution”, and as noted earlier, this is often referred to as “spatial resolution”.

156. The second usage of the term denotes the information content of an image or video picture. This is a measure of the spatial frequencies in the content, and can be measured on a display optically by the use of a spectrometer, or visually by displaying test patterns. It must be noted this is not simply determined by the pixel resolution. If one displays an image or video having the legacy NTSC number of lines on a 4K TV by “blowing up” the image to fill the screen (the patent specification uses the term “enlargement”, see below), the information resolution (spatial frequency content) of the displayed image or video will not be increased. In fact, the image will appear blurred. This usage could be called “information resolution”, examples of such usage can be found in claim 1 of the ’448 Patent, see below.

157. The patent specification uses the term “resolution” in both ways, without clarifying which meaning is being used. The problem is compounded by the patents use of the terms “enlargement” and the term “resolution conversion” again without clarifying what is meant by the terms with respect to “pixel resolution” and “information resolution”. *See below* for annotated claim 1 of the ’448 Patent.



2. Combination of Processes; Claims 1-3

158. Claims 1-3 include aspects that represent a conventional video encoding process. Video encoders, well known in the art, also perform decoding in order to create decoded pictures for use in coding successive pictures. *See* Technical Background section.

“a first combination of processes for **an encoding and a decoding** to create a first sequence of encoded bits and a set of decoded pictures”

159. Video encoders also perform prediction and encoding, i.e., entropy encoding.

“a third combination of processes for **a prediction and an encoding** to create a third sequence of encoded bits,”

3. Enlargement, Enlarger; Claims 1-3

160. Interpolating new pixels at sub-pixel locations of an original pixel array, resulting in an “enlarged” pixel array is described in various aspects of the ’448 Patent. *See e.g.*, Ex.1001 [’448 Patent], Abstract, 3:20-34, Fig. 4(a), 1032. This is also detailed in the Technical Background section above.

161. This enlargement process is just one step in the process of performing superresolution processing. *See e.g.*, Ex.1001 [’448 Patent], Fig. 4(b), 22:61-63; 23:4-27.

4. Resolution Conversion, Converter, Converted; Claims 1-3

162. As used in the patent specification, this term may mean down-sampling or up-sampling. If down-sampling is performed correctly, the array of pixels comprising an image or video picture is spatially filtered to band-limit the information, and then a subsampled array can be extracted from the array of pixels.

163. As discussed above, up-sampling in the patent specification may mean the same as “enlargement”. As noted in the Technical Background, performing resolution conversion by merely up-sampling a picture does not increase the information resolution of the picture.

164. Down-sampling is one step in the process of performing superresolution processing in the case in which the superresolution processing is performed to restore the spatial frequency content of an image within the theoretical limits of the pixel array dimensions. (See “the first way” discussion of superresolution in the Technical Background section above). In the patent specification, the input image has been “enlarged” by interpolating new pixels at sub-pixel locations for use in superresolution processing. Subsequently, the image is returned to its original “standard” resolution by down-sampling.

165. In the '448 Patent, resolution conversion may be disclosed in conjunction with superresolution or without.

5. Superresolution enlargement ... to create a set of super-resolution enlarged [decoded] pictures with a resolution higher than the standard resolution; Claims 1-3

166. The superresolution process, well-known in the art, in which input and output pixel arrays have the same size, comprises a series of operations that the patent specification articulates in the specification and claims in a particular way.

167. In the art, the superresolution process performed “in the first way” above comprises the following steps:

- a. Establish a set of input pictures that include the same content
- b. Perform sub-pixel interpolation on each picture to provide a set of pictures having a larger size array of interpolated pixels

- c. Register the set of pixels by translating each of the larger pixel arrays such that the objects imaged in each picture are aligned
- d. Average the aligned pixels in each picture to provide a single output picture
- e. Downsample the output picture to the original pixel array dimensions

168. In the art, the superresolution process performed “in the second way” in which the processed image possesses a pixel array having larger dimensions than the input image, the last step is not performed. (See Technical Background above).

169. In the patent specification, superresolution enlargement process is illustrated in Figure 4 and described at Ex.1001 [’448 Patent], 22:54-24:67. The resolution conversion process is illustrated in Figure 6 and described at Ex.1001 [’448 Patent], 25:1-27:17.

**6. Sequence, sub-sequence (of pictures), Set (of pictures);
Claims 1-3**

170. The patent specification uses the term “sequence” in an overloaded manner. First, it may mean the original sequence of video pictures input to an encoder and the encoded bitstream representing that sequence of pictures in a one-to-one mapping of each picture in the sequence.

171. Second, it may mean the subsequence of pictures from such a sequence of frames that are used in the multi-picture process to implement superresolution processing. Such a “sequence” of input pictures is used to produce just one of the sequence of output pictures, i.e., a many-to-one mapping.

172. Third, it may mean a sequence of pictures output from superresolution processor that resulted from the use of multiple subsequences of pictures from a sequence of input pictures that are used as reference pictures to generate encoded bitstreams representing a sequence of pictures, i.e., a many-to-many mapping.

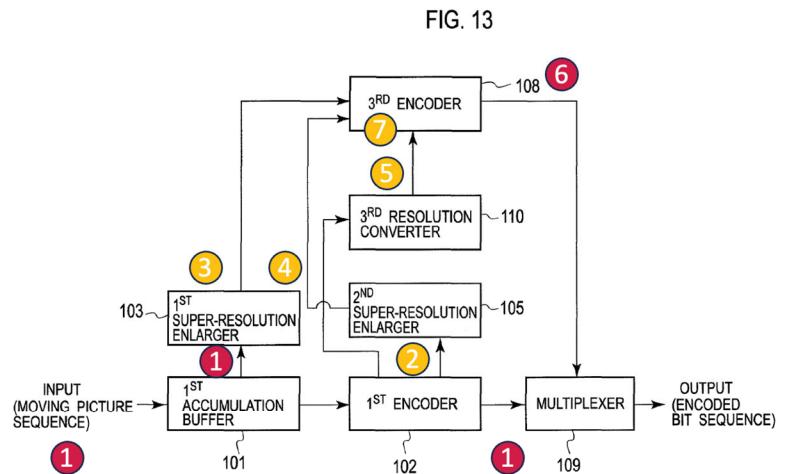
173. Fourth, it may mean a sequence of encoded bits of picture information or it may mean a sequence of bits of encoded reference picture selection information.

174. Likewise, the term “set” is used in an overloaded manner, see the illustration of claim 1 with Figure 13 below.

175. In particular, in the art a subsequence of a sequence of original pictures may be used to generate a single picture by superresolution enlargement and resolution conversion, i.e., a many-to-one process. In claim 1 however, the sequence of original pictures generates a set of superresolution target pictures, together with a set of decoded pictures from the first encoder that are used as a set of reference pictures for encoding the target pictures and a set of decoded superresolution pictures, also used as reference pictures for encoding the set of target pictures. The patent does not specify whether the “set” of such pictures are generated from the

same “sequence” of original (input) pictures, or whether different sequences of input pictures are needed to generate the various sets of pictures.

1. A moving picture encoding system that makes an encoding of a sequence of moving pictures with a resolution higher than a standard resolution using moving pictures contents which include a sequence of moving pictures with the standard resolution and do not include a sequence of moving pictures with a resolution higher than the standard resolution, the moving picture encoding system comprising:
 - 1 a first encoder configured to work on a sequence of moving pictures with a standard resolution to implement a first combination of processes for an encoding and a decoding to create a first sequence of encoded bits and a set of decoded pictures with the standard resolution;
 - 2 a first super-resolution enlarger configured to work on the sequence of moving pictures with the standard resolution to implement a process for a first super-resolution enlargement to create a set of super-resolution enlarged pictures with a resolution higher than the standard resolution;
 - 3 a second super-resolution enlarger configured to acquire the set of decoded pictures from the first encoder to implement thereon a process for a second super-resolution enlargement to create a set of super-resolution enlarged decoded pictures with a resolution higher than the standard resolution;
 - 4 a third resolution converter configured to acquire the set of decoded pictures from the first encoder to implement thereon a process for a third resolution conversion to create a set of resolution converted enlarged decoded pictures with a resolution higher than the standard resolution; and
 - 5 a third encoder configured to have the set of super-resolution enlarged pictures from the first super-resolution enlarger as a set of encoding target pictures, employing the set of super-resolution enlarged decoded pictures from the second super-resolution enlarger and the set of resolution converted enlarged decoded pictures from the third resolution converter as sets of reference pictures, to implement thereon a third combination of processes for a prediction and an encoding to create a third sequence of encoded bits,
 - 6 wherein a spatial resolution of the set of super-resolution enlarged pictures, that of the set of super-resolution enlarged decoded pictures, and that of the set of resolution converted enlarged decoded pictures are made equal; and
 - 7 wherein the third encoder controls a selection of a set of reference pictures and creates a set of data on the selection of the set of reference pictures to identify a selected set of reference pictures during the process for prediction of the third combination.



7. Encoding; Claims 1-3

176. This is an overloaded term in the art. The first usage of the term denotes the whole process of encoding video, i.e., encoding sequences of pictures. The second usage refers to the process within that whole process, of entropy encoding information to create an encoded bitstream. For example, the patent claims include encoding the information of the pixels in the video pictures and encoding the

information for selection of reference pictures and the information for encoding parameters.

B. The patent specification

1. Field of Art

177. The stated field of art is systems methods and programs for encoding, decoding and reencoding moving pictures, “adapted to work on moving picture sequences to implement processes such as those for superresolution enlargement.” Ex.1001 (’448 Patent) Title, 1:23-30.

2. Background - Problem Statement, Admitted Prior Art

178. The patent discloses a two or more layer hierarchical video coding system. The patent admits there is prior art in the MPEG standards and the SVC - scalable video coding – of the MPEG-4 standard. A stated problem with the existing art is stated to be: “failures to allot a sufficient code rate would cause deficiencies” Ex.1001 (’448 Patent), 2:12-16 and “failures to allot a sufficient code rate would degrade image qualities of predictive pictures or predictive blocks.” Ex.1001 (’448 Patent), 2:27-33.

179. A second stated problem is: “in such the [sic] hierarchical encoding, even if the encoding process implemented was [sic] reversible, input moving pictures would undergo a band limitation to information on their spatial frequency components in the spatial direction that can be represented by the standard

resolution,” Ex.1001 (’448 Patent), 2:42-59; and that while inter-layer prediction in hierarchical coding allows estimates to be made of new information on spatial frequency components, this is “failing to encode and transmit more information on spatial frequency components than the information on spatial frequency components that can be represented by the standard resolution.” Ex. 1001 (’448 Patent), 2:42-59. Therefore an object of the techniques described in the ’448 Patent is “to encode and transmit more information on spatial frequency components than information on spatial frequency components that can be represented by a resolution of input moving pictures.”

180. The patent also extensively discloses admitted prior art for superresolution. *See* Ex.1001 (’448 Patent), 22:6-50. The patent does not claim invention of any superresolution technique – it merely uses any such technique as a filtering process in conjunction with coding video and the selection between filtered or unfiltered data for employing reference pictures for motion compensation: “According to the present embodiment, the process for the prescribed super-resolution enlargement may for instance be an application of any such super-resolution process as described above.” Ex. 1001 (’448 Patent), 22:51-54.

3. Summary of the disclosed subject matter

181. The technologies employed by the Patent are not described in a clear manner in the specification, which obfuscates the simplicity of the technologies and

that they are in fact well-known in the art. Several of the known concepts implicated by the '448 Patent include:

- Resolution; standard resolution
- Enlargement, enlarger, enlarged; superresolution enlargement
- Conversion, converter; resolution converter
- Extended, extension
- Subsequence; two or more pictures

182. As discussed above, the '448 Patent uses the term “resolution” to refer to the information content of the image and spatial frequencies of the image in some instances, i.e., the “true” resolution of the image as perceived when presented to a human viewer, while in other instances the term is used to refer to the number of pixels in the image.

183. The '448 Patent specification uses the term enlargement to refer to an upsampling of the number of pixels representing the image information to a larger number. The specification then refers to the enlarged image as “higher resolution”. As explained above, such a use of the term is found in the art, but it does not mean the information content of the image is equal to the theoretical maximum spatial frequency content of that number of “higher resolution” pixels. The specification uses the term enlargement to refer to part of a sequence of operations that also include “resolution conversion.” As discussed below, although the '448 Patent

specification separates these two operations and identifies them separately in figures and descriptions, they can in fact be internal steps in the disclosed process of superresolution by registering multiple images.

184. The '448 Patent specification describes conversions in some contexts as a downsampling of the number of pixels representing the image information to a smaller number. The patent uses the term “resolution conversion”, but like the case above with enlargement, this operation does not necessarily change information content of the image, although it reduces the number of pixels to a smaller number.

... a first super-resolution enlarger configured to work on the subsequence of the sequence of moving pictures with the standard resolution to implement a process for a first super-resolution enlargement to create a set of super-resolution enlarged pictures with a resolution higher than the standard resolution, a first resolution converter configured to work on the set of super-resolution enlarged pictures to implement a process for a first resolution conversion to create a set of super-resolution enlarged and converted pictures with the standard resolution, ...

Ex.1001 ('448 Patent) 3:20-29.

185. The '448 Patent specification also discusses “extending” the frequencies of the image data in certain parts of the specification. The context of this use includes in a first case increasing the range of spatial frequencies within an image having the same number of pixels for the input image and the output image, but as noted earlier, may also in a second case include increasing the range of spatial frequencies beyond those covered by the number of pixels in the input image.

186. In the art for the second case, processes for enhancing a single image using superresolution generally involve 1) accepting as inputs, multiple images that are imaging the same scene but with different spatial offsets (spatial translation), 2) registering the pixel arrays of the multiple image using sub-pixel accuracy, 3) filtering, decimating and outputting the enhanced image containing higher spatial frequency information with an increased number of pixels from the input image.

187. As discussed in the Technical Background, one approach to registering pixel arrays with sub-pixel accuracy is to first interpolate the pixel arrays into arrays that are larger by the inverse of the sub-pixel resolution, then spatially translate those arrays to align the images of the objects being imaged. However, this is a brute force approach that is very wasteful in both computation and memory resources. For example, if the registration process is performed with $1/16^{\text{th}}$ pixel accuracy and the enhanced output images are increased in size by a factor of two horizontally and vertically, then the brute force approach would require processing and storage of 256-times the number of input pixels while the output would produce only 4-times the number of pixels.

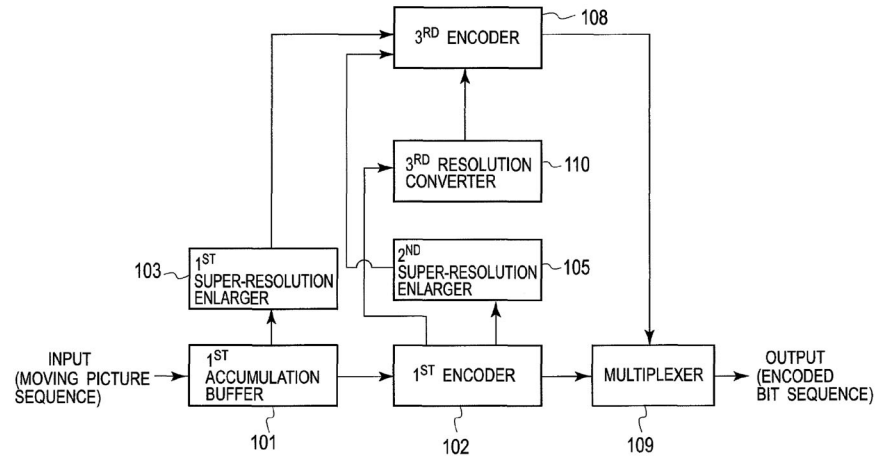
188. The '448 Patent specification discloses the second case in the fourth embodiment with respect to Fig. 13:

The first accumulation buffer 101 is configured for equivalent functions to the embodiment shown in FIG. 1, and redundant description is omitted. The first encoder 102 is configured for

equivalent functions to the embodiment shown in FIG. 1. In the fourth embodiment, there is a set of decoded pictures created to supply to the second super-resolution enlarger 105 and the third resolution converter 110. The first super-resolution enlarger 103 is configured for equivalent functions to the embodiment shown in FIG. 1. According to the fourth embodiment, the first super-resolution enlarger 103 is configured for functions to create superresolution enlarged pictures with a resolution higher than a standard resolution, to supply to the third encoder 108. The second super-resolution enlarger 105 is configured for equivalent functions to the embodiment shown in FIG. 1. According to the fourth embodiment, the second super-resolution enlarger 105 is configured for functions to create superresolution enlarged decoded pictures with a resolution higher than the standard resolution, to supply to the third encoder 108. The third resolution converter 110 is configured for functions to acquire decoded pictures with the standard resolution from the first encoder 102, and implement a process for a prescribed resolution conversion on acquired decoded pictures with the standard resolution to create resolution conversion enlarged decoded pictures being high resolution enlarged decoded pictures with a resolution equal to the spatial resolution of super-resolution enlarged pictures created at the first super-resolution enlarger 103 and that of super-resolution enlarged decoded pictures created at the second superresolution enlarger 105, to supply to the third encoder 108.

Ex.1001 ('448 Patent) 34:5-36.

FIG. 13



189. The '448 Patent specification discloses variations of configuration in Figs. 16 and 17, in which only one of the paths from the 1st Encoder to the 3rd Encoder is depicted.

FIG. 16

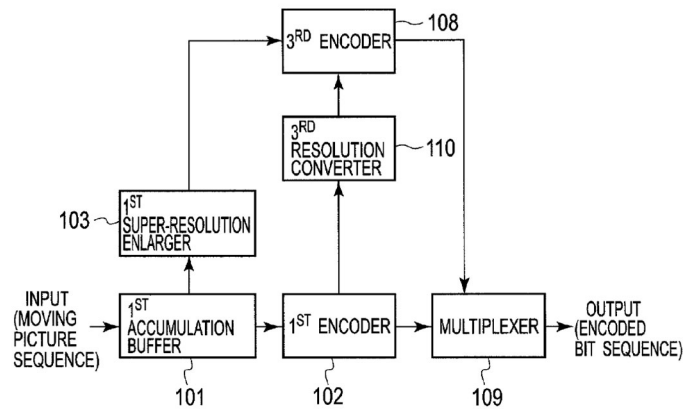
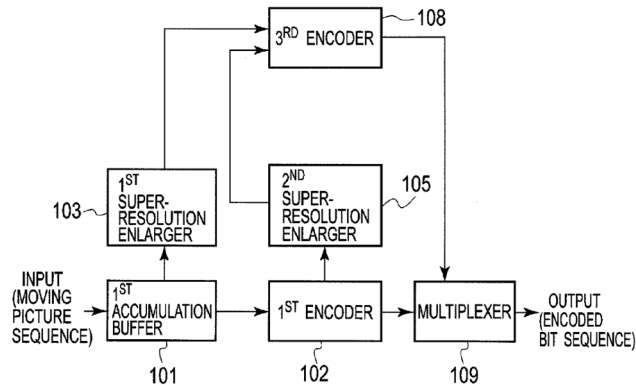


FIG. 17



190. In each case, the target pictures encoded by the 3rd encoder are a higher spatial resolution (pixel resolution) than the Input sequences.

191. The “equivalent functions to the embodiment shown in Fig. 1” for “a first accumulation buffer 101, a first encoder 102, a first superresolution enlarger 103 and a second resolution converter 105, refer to the specification at 13:23-14:49.

However, all that is disclosed for the first accumulation buffer 101 is that it:

“is configured for functions to work on moving pictures input as an input, to **accumulate two or more pictures as necessary to implement later described processes** for a prescribed pre-super-resolution encoding and a first super-resolution enlargement, and supply the first encoder 102, the first super-resolution enlarger 103, and the like **with pictures they require.**”

Ex.1001 ('448 Patent) 13:29-35 (emphasis added).

192. Both the phrases “as necessary” and “they require” are somewhat unclear. For example, these phrases imply that it does not matter how one performs super-resolution enlargement or encoding.

193. All that is disclosed for the first encoder 102 is that it:

“is configured for functions to **acquire** from the first accumulation buffer 101 **a set of pictures** with a standard resolution **required for a process for a prescribed encoding**, and work on pictures acquired with the standard resolution to implement the process for the prescribed encoding (referred herein to as a process for a pre-super-resolution encoding), to create a first sequence of encoded bits, to supply the first sequence of encoded bits thus created to the multiplexer 109. The first encoder 102 is configured for functions to implement **a process for a prescribed decoding**, to supply decoded pictures to the second super-resolution enlarger 105 and the second encoder 107. The first encoder 102 may well be configured for functions to **temporarily store decoded pictures**. There may be a configuration to implement the process for the prescribed decoding as a decoding process on the first sequence of encoded bits as created to create decoded pictures. There may be configuration for use of a process for a local decoding constituting part of the process for the prescribed encoding to employ results of the local decoding to create pictures decoded therefrom.”

Ex.1001 ('448 Patent) 13:36-55 (emphasis added).

194. The phrases “a set of pictures ... required for a process for a prescribed encoding.”, “a process for a prescribed decoding” and “temporarily store decoded pictures” are also somewhat unclear. For example, these phrases imply that it does not matter how one performs encoding or decoding. In particular, no guidance is provided for how many pictures must be in the “set of pictures” or temporarily stored.

195. What is disclosed for the first superresolution enlarger 103 is that:

“The first super-resolution enlarger 103 is configured for functions to **acquire** from the first accumulation buffer 101 **two or more pictures** with the standard resolution, **as necessary** for a process for a **prescribed super-resolution enlargement**, and work on pictures acquired with the standard resolution to implement the process for the prescribed superresolution enlargement, to create super-resolution enlarged pictures with a resolution higher than the standard resolution, and supply thus created super-resolution enlarged pictures to the first resolution converter 104. There may be configuration with a function to have a ratio of enlargement set up upon implementation of the process for the prescribed super-resolution enlargement as preset information on a prescribed enlargement ratio, or as acquired information on enlargement ratio such as set up by an external user, and work on the basis of acquired information on enlargement ratio to establish an enlargement ratio to implement the process for the prescribed super-resolution enlargement. More preferably, to provide information on the prescribed enlargement ratio to a decoding system end, there should be use of a configuration to implement an unshown process for a prescribed entropy encoding, to create a bit sequence of information on enlargement ratio, to supply to the multiplexer 109, to transmit to that end. In this regard, there should be use of similar configuration for the second super-resolution enlarger 105 also. The first superresolution enlarger 103 and the second super-resolution enlarger 105 have their configurations and actions, of which detailed examples will be described with reference to FIG. 4.”

Ex.1001 ('448 Patent) 13:56-14:16 (emphasis added).

196. Both the phrases “two or more pictures” and “as necessary” are somewhat unclear. For example, these phrases imply that it does not matter how one performs super-resolution enlargement.

197. Figure 4 is disclosed as follows:

“Part (a) of FIG. 4 is a block diagram showing a detailed configuration of the first super-resolution enlarger 103 in the first embodiment according to the present invention, Part (b) of FIG. 4, a combination of conceptual diagrams illustrating a process for super-resolution enlargement, ... In Part (a) of FIG. 4, according to the first embodiment, the first super-resolution enlarger 103 includes a positioner 1031, an interpolator 1032, an estimated picture creator 1033, and a repetition determiner 1034. It is noted that the second the super-resolution enlarger 105 has a similar configuration.”

Ex.1001 ('448 Patent) 22:59-23:3.

198. The disclosure continues:

The positioner 1031 is configured to work in a processing for super-resolution enlargement, to acquire combination of a base picture constituting a base in therein and one or more observation pictures, such as those illustrated in Section (I) in Part (b) of FIG. 4. It is now assumed to acquire such a combination of base picture, observation picture 1, and observation picture 2, as illustrated in Section (I) in Part (b) of FIG. 4. As illustrated in Section (II) in Part (b) of FIG. 4, the positioner 1031 is configured for functions to make a registration or positioning to pixel positions of a resolution higher than resolutions of observation pictures 1 and 2 and base picture thus acquired, for **an interval-unequal sampling** to create **nonhomogeneous high resolution pictures**. Further, it is configured for a function to supply created nonhomogeneous high resolution pictures to the interpolator 1032 and the repetition determiner 1034. For the creation of nonhomogeneous high resolution pictures, preferably, as illustrated in Section (II) in Part (b) of FIG. 4, observation pictures should be each positioned relative to pixels of the base picture, to arrange in place to have a highest correlation with the base picture.

Ex.1001 ('448 Patent) 23:4-24 (emphasis added).

199. The specification does not clearly explain how the interpolation would be performed and does not always clearly describe the buffers needed to hold the enlarged pixel arrays. The specification also does not describe the process to register the enlarged arrays of pixels of all members of the “subsequence” or “two or more pictures” that may result in higher spatial frequency content. In particular, the specification does not explain what is meant by “interval-unequal sampling” or “nonhomogeneous high-resolution pictures”.

200. The specification at 22:7-50 lists a number of prior art references concerning superresolution. These include the Park reference (misstated as the “Sung” reference – see footnote above) and discloses that how superresolution is performed is irrelevant:

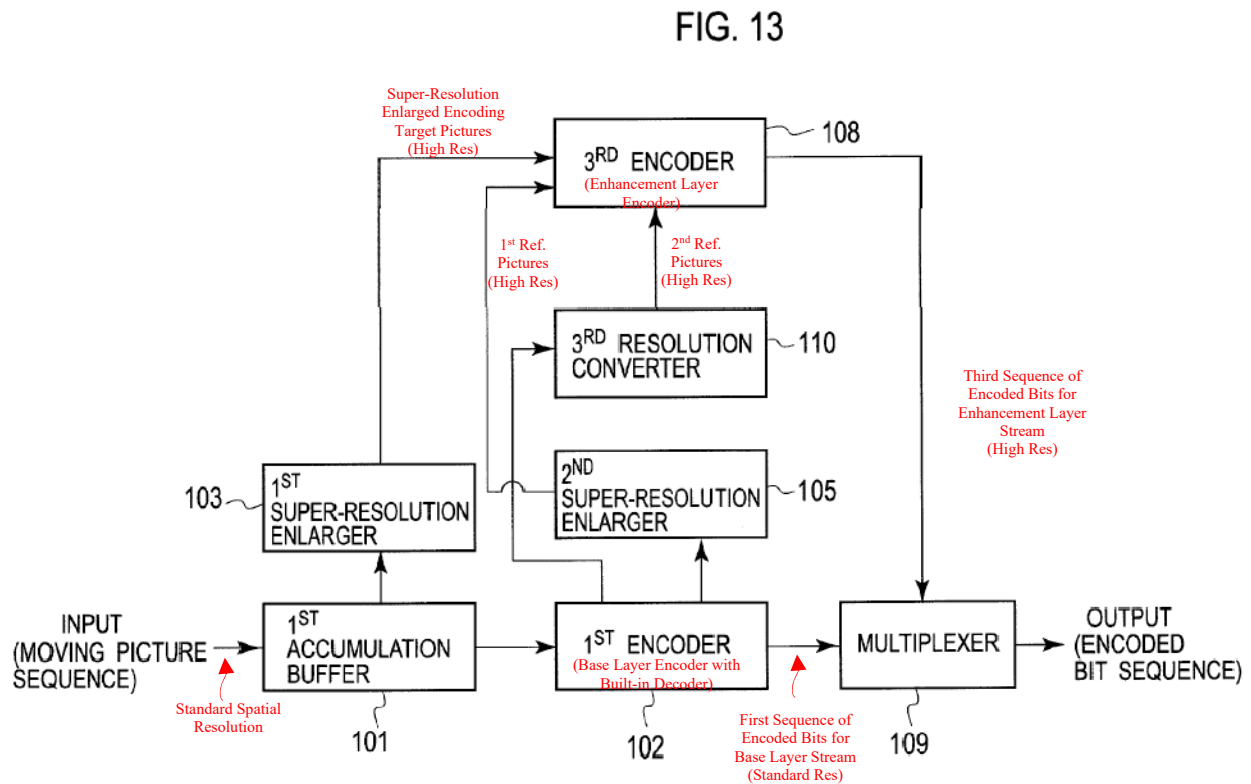
According to the present embodiment, the process for the prescribed super-resolution enlargement may **for instance** be an application of **any** such super-resolution process as described above.

Ex.1001 ('448 Patent) 22:51-54 (emphasis added).

201. I note a strong resemblance between the Park reference and Figure 4.

202. The '448 Patent describes a layered video coding system that employs conventional super-resolution algorithms with an objective to enhance the quality of the original pictures and the base-layer reference pictures in an enhancement layer.

203. Figure 13 of the '448 Patent is instructive, as its configuration closely parallels limitations recited in the challenged claims:



204. As shown above, the encoding apparatus passes pictures from the input video through an accumulation buffer 101 to each of two encoding paths. Ex.1001 ['448 Patent], 34:5-7, 13:20-17:16. In the first path, a base version of the input video is encoded by the first encoder 102 to produce a first coded bitstream for the base version of the video. *Id.*, 34:8-12, 13:29-13:55. In the second path, a super-resolution enhanced version of the input video is encoded by the third encoder 108 to produce a second coded bitstream for the super-resolution enhanced version of

the input video. *Id.*, 34:46-61. The system generates the super-resolution enhanced version of the input video using a first super-resolution enlarger 103 to increase the spatial resolution and information content (e.g., quality) of pictures in the input video. *Id.*, 34:13-18, 13:36-55. For each input video picture that is to be subjected to super-resolution processing, the first accumulation buffer 101 (which is not recited in the challenged claims) accumulates subsequences of pictures from the neighborhood of the input video picture that is to be super-resolution enhanced, in which the pictures in the subsequences are registered with sub-pixel accuracy to provide a corresponding super-resolution image. *Id.*, 4:27-42.

205. To remove redundancies in information coded between the base- and enhancement-layer streams, the third encoder 108 implements inter-layer predictive processes to predict pictures in the enhancement-layer stream from decoded pictures in the base-layer stream. Ex.1001 ['448 Patent], 34:46-61. Decoded pictures from the base-layer stream can also be enhanced before being used as reference by the third encoder 108 by subjecting base-layer pictures that have been encoded and then decoded by the first encoder 102 to processing by the second super-resolution enlarger 105 to create super-resolution enhanced base layer pictures. Alternatively, the decoded pictures may be subjected to processing by the third resolution converter 110 to create resolution enlarged base layer pictures. *Id.*, 34:19-45. The choice between super-resolution enlargement or resolution conversion enlargement is

dependent on which technique will be optimal, considering many factors (as discussed in the Technical Background). The third encoder 108 may select either super-resolution enhanced base layer pictures or resolution converted enlarged base layer pictures as reference pictures for encoding of the enhancement layer video stream (or neither). Ex.1001 ['448 Patent], 35:33-58. This selection of super-resolution enhanced base layer pictures or resolution converted enlarged base layer pictures as reference pictures may be on an entire picture basis, or may be on the basis of a region of a picture (e.g., a block). The base and enhancement layer streams are then multiplexed (109) into a single output signal, together with a signal indicating which base-layer picture was selected, and the multiplexed signal is transmitted to a decoding apparatus. *Id.*, 35:49-36:44, 29:30-35, 58:59-59:16.

4. Comments

206. The specification repeatedly refers to “sequences” of images and “sets” of images without identifying whether the reference is to one of the images from the sequence or to all the images in the sequence, nor to whether the member of a set was the result of processing one image from sequence, or from all the images in the sequence.

207. The '448 Patent comprises a set of very well-known image and video processing tools, including the following:

- Video coding

- The use of multiple reference frames for motion compensated prediction
- Selection of reference frames on a picture or “region” (e.g., block) basis
- Enhanced (loop filtered) reference pictures for motion compensated prediction
- Scalable video coding
 - Selection of inter-layer reference data
 - Selection on a picture or “region” (e.g., block) basis
- The use of superresolution for preprocessing

208. First invented in 1974, motion compensated prediction was incorporated (as a selectable option) in the H.261 standard.

209. Beginning in the 1980s with H.261, video coding standards with motion compensated prediction, employed the use of optionally-enabled loop filters in which on a frame level or a block level (“region”), the reference picture could be enhanced by filtering.

210. Beginning in the 1990s with H.263+, video coding standards employed the use of motion compensated prediction in which on a picture level or a block level (“region”), the reference picture could be selected from one of a number of pictures. As discussed in the background section, there is no way to mathematically define which reference picture is optimum. An encoder determines such empirically, based

on the local picture data characteristic and the actual minimum number of generated bits and signals to the decoder which frame to use.

211. The '448 Patent specification employs these well-known methods and proposes for example that the reference pictures might comprise an unprocessed reference picture or a picture subjected to the superresolution processing. Due to the random nature of video data, either one may be optimum on a case by case basis.

212. The scalable video coding standards provide for prediction of a picture in an enhanced layer by temporal prediction within the same layer and/or prediction from a contemporaneous picture at a lower layer. The '448 Patent specification proposes the same thing.

213. The '448 Patent specification proposes using a filtered reference frame in which the filtering process is the well-known superresolution process.

214. Regarding superresolution, for a single picture k , a subsequence of pictures, say $k-n$ to $k+m$ is used. For each picture k , a spatial registration process is performed, in which at a sub-pixel accuracy the data in the picture is spatially translated to align the image of the objects in the scene. The aligned images are then combined, and filtered to produce a single output picture k' that has the same number of pixels as the input picture k , but may have a higher spatial frequency content.

C. Prosecution History of the '448 Patent

215. I have reviewed the Prosecution History of the '448 Patent. *See* Ex.1002. The application (U.S. Serial No. 14/692,138) resulting in the '448 Patent was filed on May 28, 2009, by Satoru Sakazume. The '448 Patent claims priority to several earlier-filed applications, with the earliest claimed priority date being May 30, 2008. *Id.*

216. During original examination, the Examiner required the applicant to narrow the application that issued as the '448 Patent to a single species of claims. Ex.1002 ('448 FH), 230-235. In response, the applicant elected a single species—"Species D" and Figure 13—and identified a limited set of claims encompassing that species. *Id.*, 237-239. In the following Office Action, the Examiner rejected claims 7 (issued claim 1), claim 8 (issued claim 2), and claim 9 (issued claim 3) as obvious over prior art combinations based on Phek¹³ (Ex.1031 [Phek]), along with additional references including Holcomb (Ex.1024 [Holcomb]) and Wan (Ex.1049 [Wan]), and explained that the remaining claims were not generic and were not included in Figure 13. Ex.1002 ('448 FH), 245-261. The applicant subsequently amended claim 7 (issued claim 1), claim 8 (issued claim 2), and claim 9 (issued claim 3), and withdrew

¹³ The file history refers to a machine translation of the Phek prior art reference by the name "Ong" rather than Phek.

the remaining claims. *Id.*, 317-344. The Examiner subsequently allowed the application. *Id.*, 352-362. The file history contains no indication that the Examiner ever considered Phek's teachings in combination with the teachings of Martins and He as discussed below in Ground 1.

VI. Summary of the Applied Prior Art

217. The following pieces of prior art render obvious the alleged invention of the '448 Patent in the combinations addressed further below (Grounds 1-2).

A. Overview of Phek

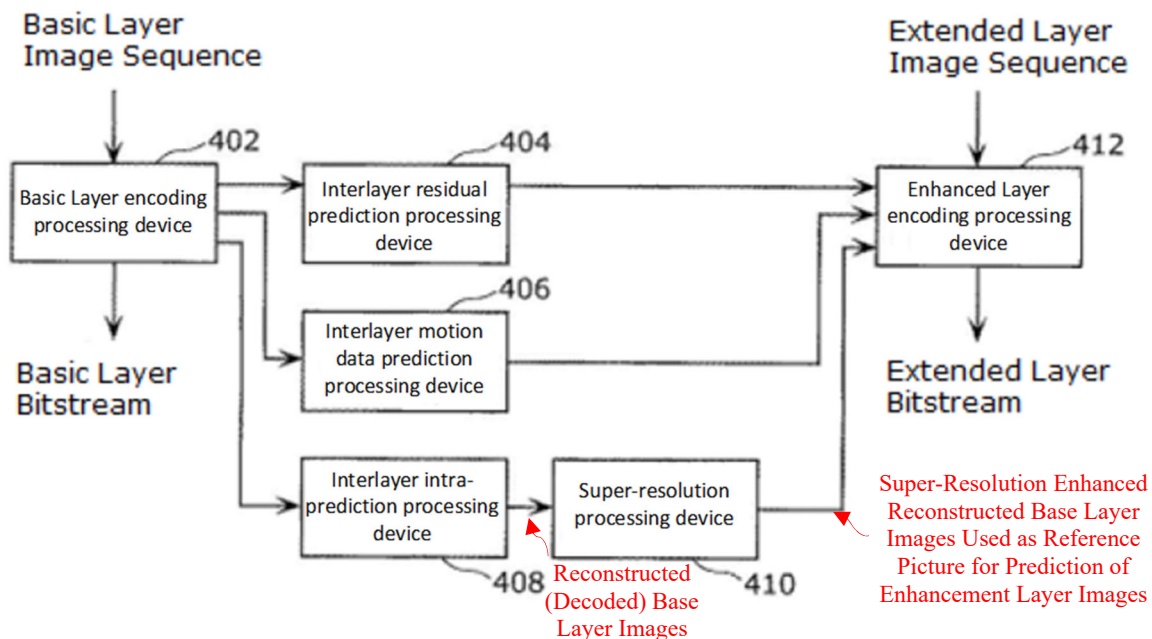
218. Phek states a problem in the Abstract: "To provide a super-resolution processing method and device that improves the resolution and quality of image sequences and that reduces grid artifacts" and provides a comprehensive solution that includes these key features:

- Generation of a superresolution image from an input image (*see e.g.*, Ex.1005 [Phek], claim 1), i.e., generate a higher resolution image from a standard resolution image (only).
- Phek's input image may be processed or unprocessed; Phek's claimed device may be a pre-processor or a post-processor, i.e., Phek's device may operate on raw input image or a decoded image. *See e.g.*, Ex.1005 [Phek], claim 10, [0002], [0010], Fig. 2.

- Phek teaches that superresolution may be used for the reference images that are used for motion compensation. *See e.g.*, Ex.1005 [Phek], Claim 11, Fig. 3, [0011].
- Phek teaches that superresolution may be used in scalable applications on decoded base images that are used for interlayer prediction as reference images. *See e.g.*, Ex.1005 [Phek], Claim 13, [0012], [0013], [0014], Figs. 4a & 4b.
- Phek discloses generation of superresolution enhanced pixels or alternatively interpolation of pixels. *See e.g.*, Ex.1005 [Phek], [0005], [0011], [0012], [0014].
- Phek provides a superresolution process description (i.e., interpolation, registration, etc.), which may use compressed video ancillary data, e.g., motion vectors. *See e.g.*, Ex.1005 [Phek], [0010], [0011], [0012], [0014], [0016], Fig. 1a.
- Phek provides a description of an interpolation process. *See e.g.*, Ex.1005 [Phek], Fig. 6a & 6b, [0018]-[0020]

219. Phek discloses a video coding system that uses “[i]nterlayer prediction ... for hierarchical encoding in order to improve encoding efficiency.” Ex.1005, [0014]. “One example of hierarchical encoding that uses interlayer prediction is ... scalable video encoding.” *Id.* Phek’s interlayer coding system includes a basic layer

and an enhanced layer, where “an intra image of the basic layer or a reconstructed image of the basic layer” is used “as an additional reference image for encoding the image of the enhanced layer.” *Id.* Moreover, “super-resolution processing ... is used to improve these additional reference images” from the basic layer. *Id.*; *see also id.*, [0011]-[0020], FIGS. 1a-1b, 3a-3b, 4a-4b, 5.



Ex.1005, FIG. 4b (modified and annotated)¹⁴.

¹⁴ For ease of reading, labels have been added in each of Phek’s figures based on the labels provided in the certified English translation. *See* Ex.1005. Phek and English translations of Phek (Ex.1005 and Ex.1031) use the terms “base” and “basic” layers

B. Overview of Martins

220. Martins is an IEEE paper titled “A Unified Approach to Restoration, Deinterlacing and Resolution Enhancement in Decoding MPEG-2 Video.” Ex.1007 [Martins], 1. Martins discloses an algorithm by which “[t]he quality and spatial resolution of video can be improved by combining multiple pictures to form a single super-resolution picture.” *Id.*, Abstract. “The superresolution pictures obtained by the algorithm are of much higher visual quality and have lower [mean-squared error] than superresolution pictures obtained by simple spatial interpolation.” *Id.* “The algorithm is based on motion-compensated spatial upsampling from multiple images and decimation to the desired format.” *Id.*

221. Martins discloses these key features:

- Post-processing of decoded MPEG video. *See e.g.*, Ex.1007 [Martins], p. 803.

(or streams) interchangeably, and further uses the terms “extended,” “enhanced,” and “enhancement” layers (or streams) interchangeably.

In my declaration, I have provided cites to and quotes disclosure from the certified translation that uses different ones of these terms, but in my declaration, I typically refer simply to “base” and “enhancement” layers (or streams) consistent with common usage of these terms in the field.

- Multiple functionalities may be provided by Martins superresolution processing – the output can be enhanced standard definition TV (SDTV), deinterlaced SDTV, and HDTV. *See e.g.*, Ex.1007 [Martins], pp. 803-804.
- The input to Martins' superresolution processing is SDTV images, i.e., sequences of moving pictures with a standard resolution alone. *See e.g.*, Ex.1007 [Martins], pp. 803-804.
- In Martins' superresolution method, multiple pictures are aligned using motion compensation, bilinear interpolation is used for sub-pixel accuracy, mapping pixels from each reference picture is used for registration. *See e.g.*, Ex.1007 [Martins], pp. 804-806, Fig. 2.
- Martins discloses the use of upsampling via interpolation as an alternative method to superresolution when the latter leads to undefined pixels, meaning the former (interpolation) is a better selection. *See e.g.*, Ex.1007 [Martins], p. 807, Fig. 3.
- Martins discloses the use of an attribute per each superresolution picture pixel for describing the selection of superresolution or interpolation.
- Martins discloses the use of superresolution on an input standard resolution sequence (SDTV) to produce a higher resolution output sequence i.e., HDTV *See* Ex.1007 [Martins], Fig. 13, p. 811.

- Martins teaches processing decoded pictures utilizing data from the encoding and decoding process to facilitate the super-resolution enhancement using motion compensation information (e.g., motion compensation vectors). *See id.*, 1, 3.

C. Overview of He

222. He discloses techniques for selective in-loop filtering in a scalable video coding system. He specifically focuses on selective in-loop deblock filtering to control deblock filtering parameters at each layer of scalable coding hierarchy. Ex.1008, [0002]-[0006], [0022], [0028], [0032]-[0037], FIGS. 1-4.

223. According to He, “[t]he deblocking filter enhances the decoded picture quality and increases the encoding efficiency by removing the block effects in the boundary of each 4x4, 8x8, and/or 16/16 block,” although “[t]he deblocking filter[] ... also consumes a significant amount of power and processing cycles.” *Id.*, [0006]. “Disabling the entire deblocking filter [] provides a significant reduction of computational complexity and power consumption, especially in the decoder of a video processing system.” *Id.*, [0028]. He thus proposes techniques for examining “the content of the video information ... to determine whether to disable the deblock filtering,” e.g., based on “a characteristic indicative of a relative amount of content complexity and/or a relative amount of motion of the video information” and “also based on power savings considerations.” *Id.*, [0028], [0037].

224. FIG. 1 (reproduced below) depicts a video encoder 100 configured to implement selective deblock filtering. Ex.1008 [He], [0023], FIG. 1. The encoder 100 performs encoding operations to encode at least a portion of a picture and outputs a coded bitstream BTS that contains “video information” for that picture. *Id.* The encoder 100 includes a frame buffer 109 that provides reference information (REF), which “includes information from one or more previously encoded and decoded frames that have been further filtered, such as by deblocking filter 131.” *Id.* URF (the unfiltered reconstructed frame) is provided to the deblocking filter 131 which, when enabled, “filters the unfiltered reconstructed frame information URF and provides filtered reconstructed information RF to an RF buffer 133.” *Id.*, [0026]. If the deblocking filter is disabled, “the unfiltered reconstructed frame information URF is passed as the reconstructed information RF in which deblock filtering is bypassed.” *Id.*

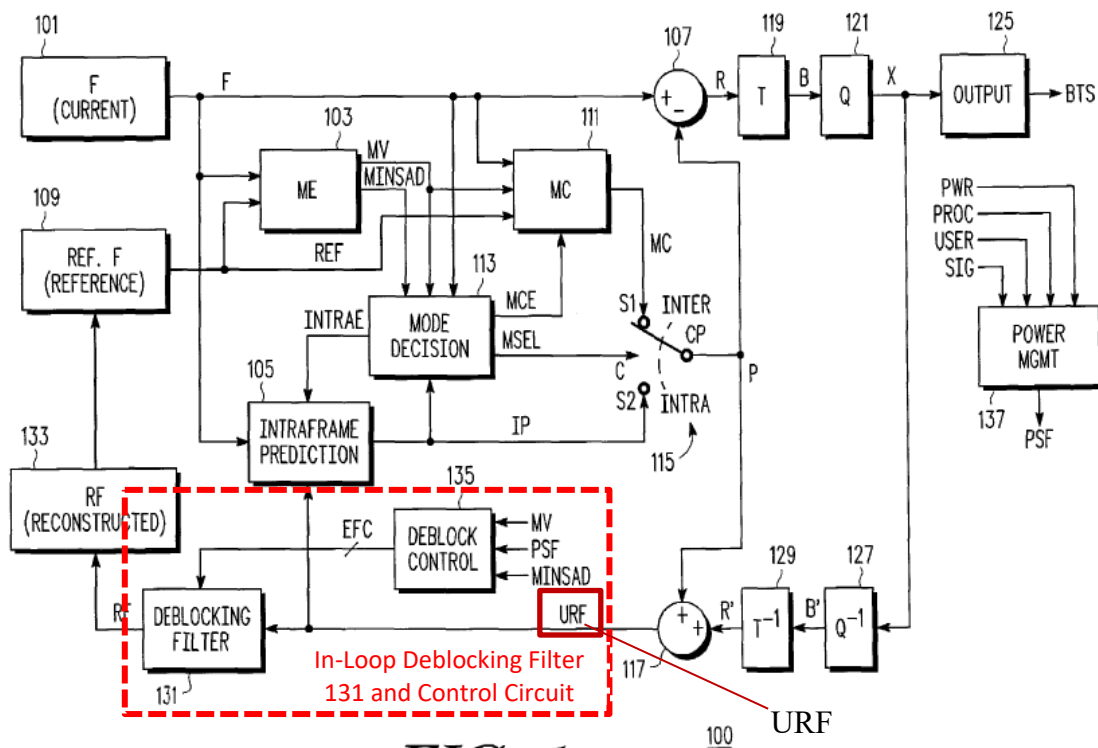


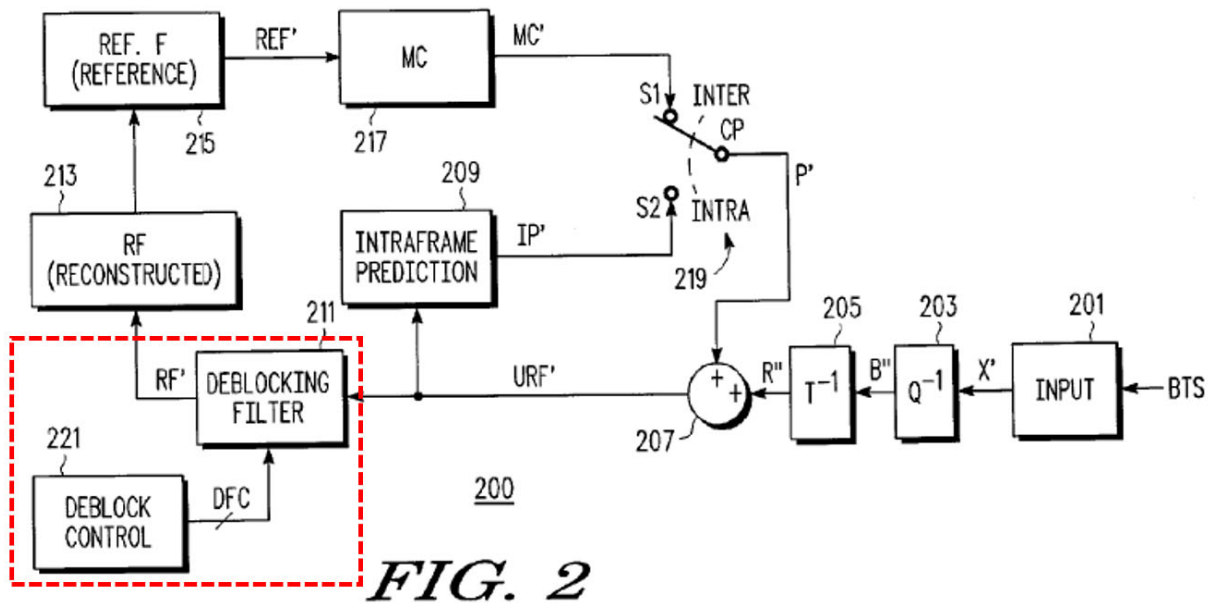
FIG. 1

Ex.1008 [He], FIG. 1 (annotated).

225. He explains that while deploying the deblocking filter can “enhance[] the decoded picture quality and increase[] the encoding efficiency,” a deblocking filter “also consumes a significant amount of power and processing cycles” and thus disabling the deblocking filter 131 significantly reduces computational complexity and power consumption. *Id.*, [0006], [0028]. For example, “[i]f the video information content exhibits a relatively low amount of motion or is relatively simple content,” “the deblocking filter 131 may be fully disabled without significant degradation in visual quality.” *Id.*

226. FIG. 2 (reproduced below) depicts a video decoder 200 configured to implement selective deblock filtering. Ex.1008, [0032], FIG. 2. The decoder 200 receives a coded bitstream (BTS) and performs prediction and reconstruction operations to decode at least a portion of a picture, thereby generating “unfiltered reconstructed frame information URF’.” *Id.*, [0032]. URF’ is provided to a deblocking filter 211, controlled by a deblock control circuit 221. *Id.*, [0033]-[0034]. Deblock control circuit 221 issues a “decoder filter control DFC signal” that selectively enables or disables the deblocking filter 211. *Id.*, [0034]. When filter 211 is disabled, deblock filtering operations are bypassed and “URF’ is passed as the ‘filtered’ reconstructed information RF’ to a reference frame buffer 213 for use as a reference picture for motion-compensated prediction processes. *Id.* When filter 211 is enabled, URF’ is filtered to generate RF’ for use as a reference picture in the motion compensated prediction processes. *Id.*

227. He further explains that “the video encoder 100 provides control information within the bitstream BTS to cause the deblock control circuit 221 of the video decoder 200 to provide the DFC signal to adaptively disable deblock filtering for a corresponding slice or frame of URF’ information being processed in the video decoder 200.” *Id.*



Ex.1008, FIG. 2 (annotated)

VII. Claim Construction

228. I have not been asked to opine on the special meanings of claim terms within the '448 Patent. I have been informed that the claim terms should be afforded their ordinary meaning in view of the specification and file history. By explaining the relevance of the prior art to the various terms and limitations in claims 1-3, I have not intended to formally construe any claim terms and I have not formed an opinion as to whether these terms are definite or supported by the specification of the '448 Patent.

VIII. Ground 1 – Claims 1-3 are Obvious under 35 U.S.C. § 103 based on Phek in view of Martins, and He

A. Overview of the Phek-Martins-He Combination

229. Like the '448 Patent, Phek discloses a layered video coding system that enhances reference picture quality in an enhancement layer by super-resolution processing. *Supra*, §III.A. While Phek's disclosure focuses on details of its super-resolution algorithm, Phek's description of the layered coding system itself (which uses the proposed super-resolution algorithm) is less detailed (though very well known in the art). Indeed, Phek acknowledges that its encoder/decoder descriptions are "simplified" and "do not have to be performed as shown." Ex.1005, [0013]. Phek's system is thus ripe for improvement.

230. Below, I address straightforward incorporations and augmentations of supporting technologies with Phek's system, which supporting technologies would have been already known by a POSITA, in light of the teachings of Martins and He¹⁵, to fill in details of Phek's invention.

1. Obvious to Generate Phek's Extended Layer Image Sequence Target Images by Super-Resolution Processing

¹⁵ This overview of the Phek- Martins-He combination is referenced and incorporated in the analysis of each claim element in Section VIII.B below.

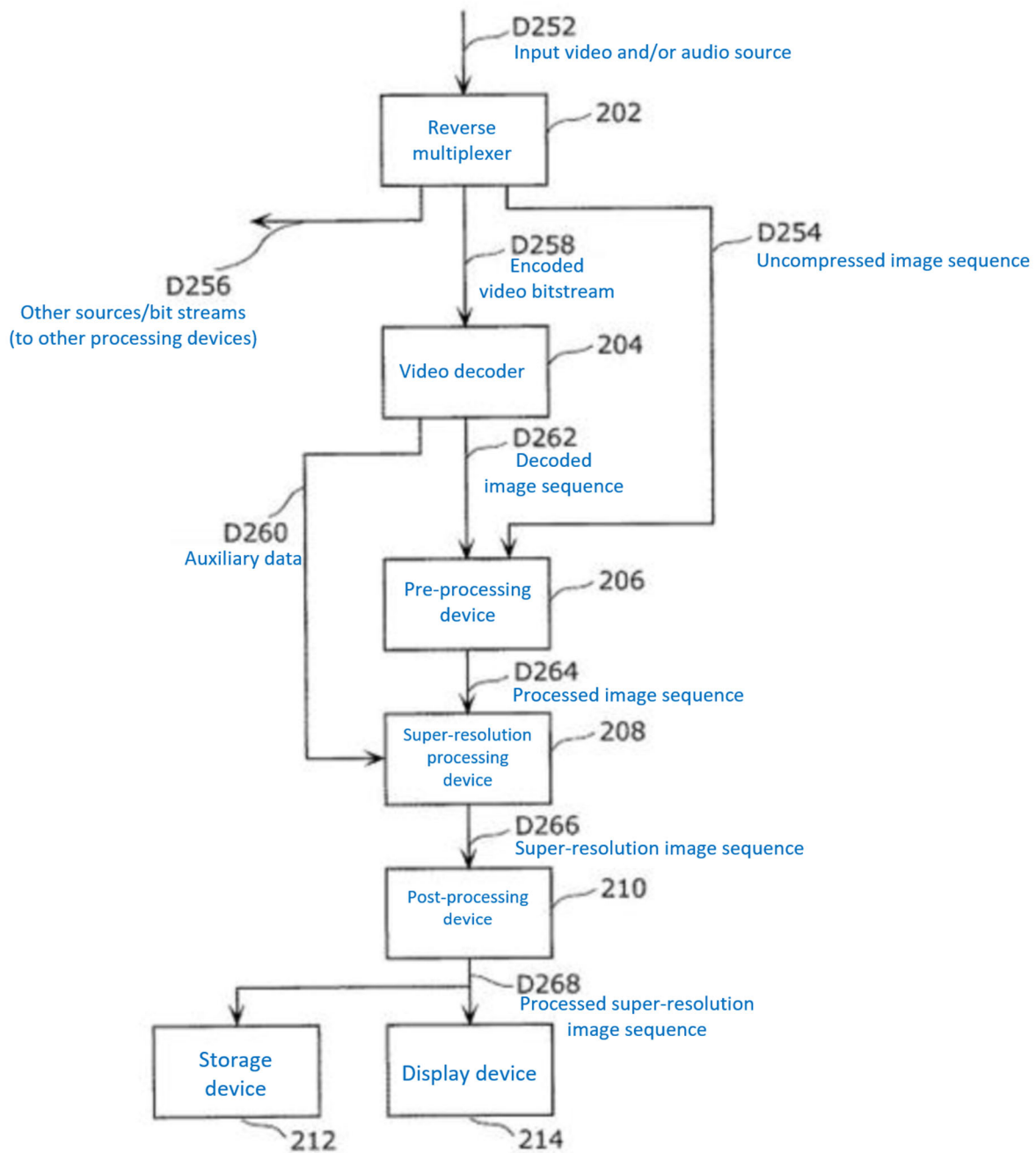
231. The challenged claims require “a first super-resolution enlarger configured to work on the sequence of moving pictures with the standard resolution to implement a process for a first super-resolution enlargement to create a set of super-resolution enlarged pictures with a resolution higher than the standard resolution.” For the reasons discussed below, this would have been obvious over the Phek-Martins-He combination, because Phek and Martins both teach such a super-resolution enlarger and it would have been obvious to use such a super-resolution enlarger in the combination.

232. As Figure 4a of Phek shows, an extended layer image sequence is provided to the enhanced layer encoding processing device 412 for encoding (i.e., the extended layer image sequence is the encoding target) to generate an extended layer bitstream. Ex.1005 [Phek], [0014], Fig. 4a. As discussed below, Phek discloses generating such a sequence directly from a base layer sequence. In the context of spatial scalability, one of the types of hierarchical coding schemes implemented by Phek’s system, this input extended layer image sequence would have been a higher spatial resolution sequence than the base layer image sequence.

233. Phek does not specify how the extended layer image sequence in Figure 4a is generated. As detailed above, however, many techniques were known for generating such an image sequence, and Phek itself includes teachings applicable to generating an extended layer image sequence. For example, a POSITA reviewing

Phek would have recognized it would have been obvious, with a reasonable expectation of success, to generate the extended layer image sequence in Figure 4a using other teachings within Phek, for example, the teachings of Figures 1a, 1b, or 2, which take as input an image sequence corresponding to raw video (such as the Base Layer image sequence) and output a super-resolution version of that same image sequence (i.e., a higher spatial resolution sequence than the base layer image sequence). For example, Phek's Figure 2 includes a pre-processing device that can perform super-resolution processing on raw data input. Ex. 1005 [Phek], [0010]. In particular, Phek's Figure 2 includes "pre-processing device 206 to generate a processed image sequence D264" which is provided to "super-resolution processing device 208" which "processes the image sequence to produce a super-resolution image sequence D266." Ex. 1005 [Phek], [0010]. The "super-resolution processing device" in turn is described in detail with reference to Figure 1a, "showing the processing devices in the super-resolution processing device" with another example in Figure 1b.

[Fig. 2]



Ex.1005 [Phek], Fig. 2 (annotated with Phek's labels)

234. As detailed above, a POSITA would have recognized that multiple prior art references at the time of the '448 Patent detailed super-resolution enhancement

techniques for generating an enhanced image sequence. To this point, additionally, or alternatively, a POSITA would have recognized that it would have been obvious, with a reasonable expectation of success, to generate the extended layer image sequence from the base layer image sequence by applying the teachings of Martins.

235. Specifically, Martins' FIG. 2 discloses "combining multiple pictures to form a single super-resolution picture" using an algorithm "based on motion-compensated spatial upsampling from multiple images." Ex.1007 [Martins], Abstract. Through upsampling, pixels of the current picture are aligned with pixels of a reference picture using "motion estimation." *Id.*, 5. The resulting information "from all the reference pictures" are combined "to form a single super-resolution picture." *Id.* Martins also discloses its super-resolution enlarger may enlarge an image using "spatial interpolation." *Id.*

236. In one possible implementation of the proposed combination of teachings, the teachings of Phek in Figure 4a are combined with the teachings of Phek in Figure 1a, for example, such that the extended layer image sequence referenced in Phek's Figure 4a is generated by applying Phek's teachings of super-resolution processing (e.g., Phek's Fig 1a) (first example).

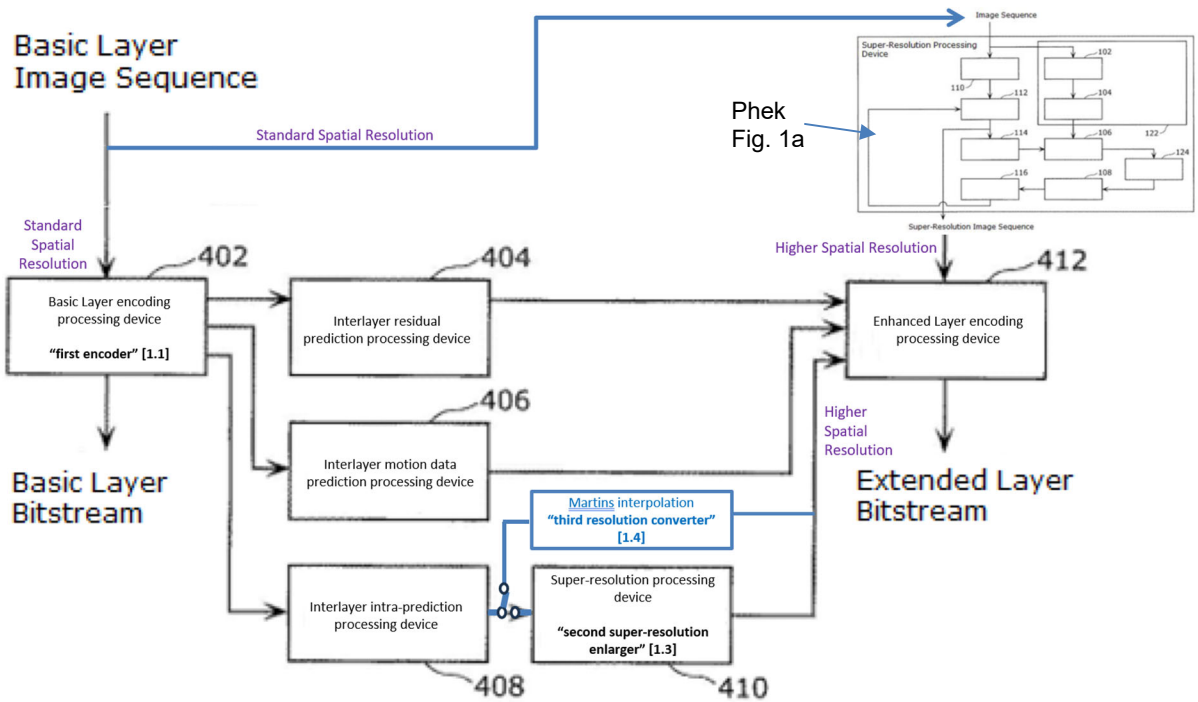


Illustration of using Phek's Fig. 1a for generating Extended Layer Image Sequence (first example)

237. For simplicity, Phek's teachings of generating a higher spatial resolution image sequence via super-resolution processing, as taught, e.g., by Phek's Fig. 1a, are represented by the illustration below. This simplified illustration is used throughout the limitation-by-limitation analysis that follows.

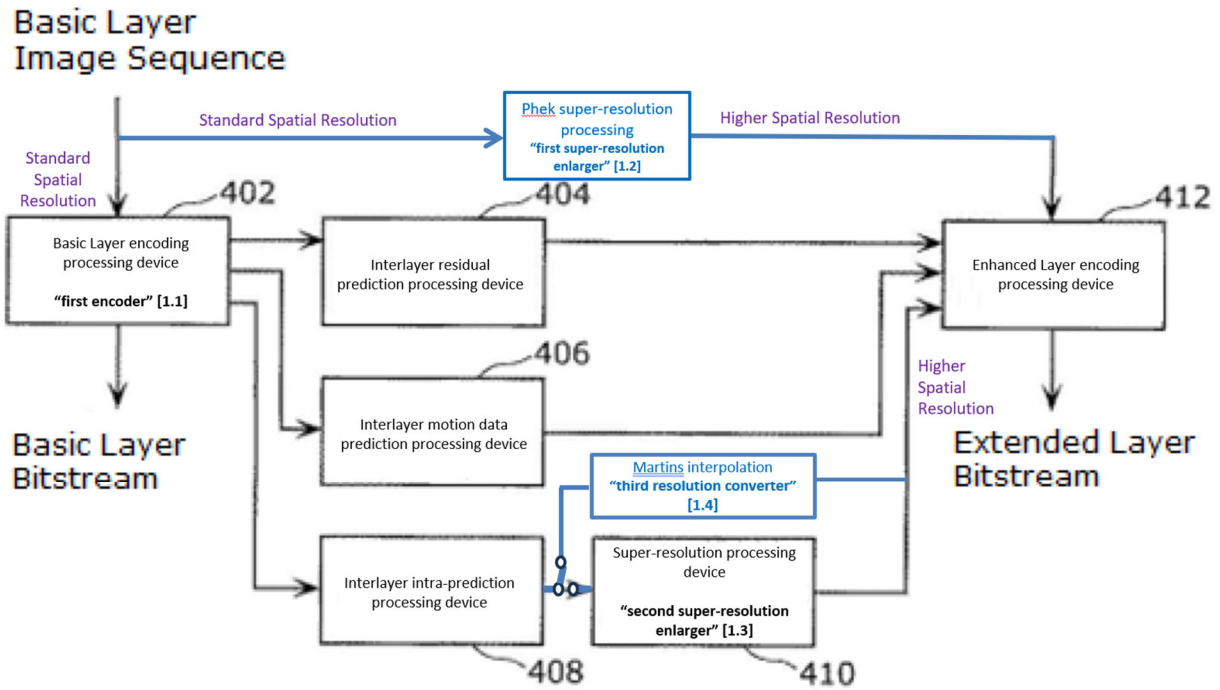


Illustration of Phek-Martins-He Combination (first example).

In another possible implementation of the proposed combination of teachings, Martins' teachings of super-resolution upsampling are combined with Phek's teachings in Figure 4a (second example), such that Martins' super-resolution upsampling is applied to Phek's base layer image sequence, as illustrated below in the second example.

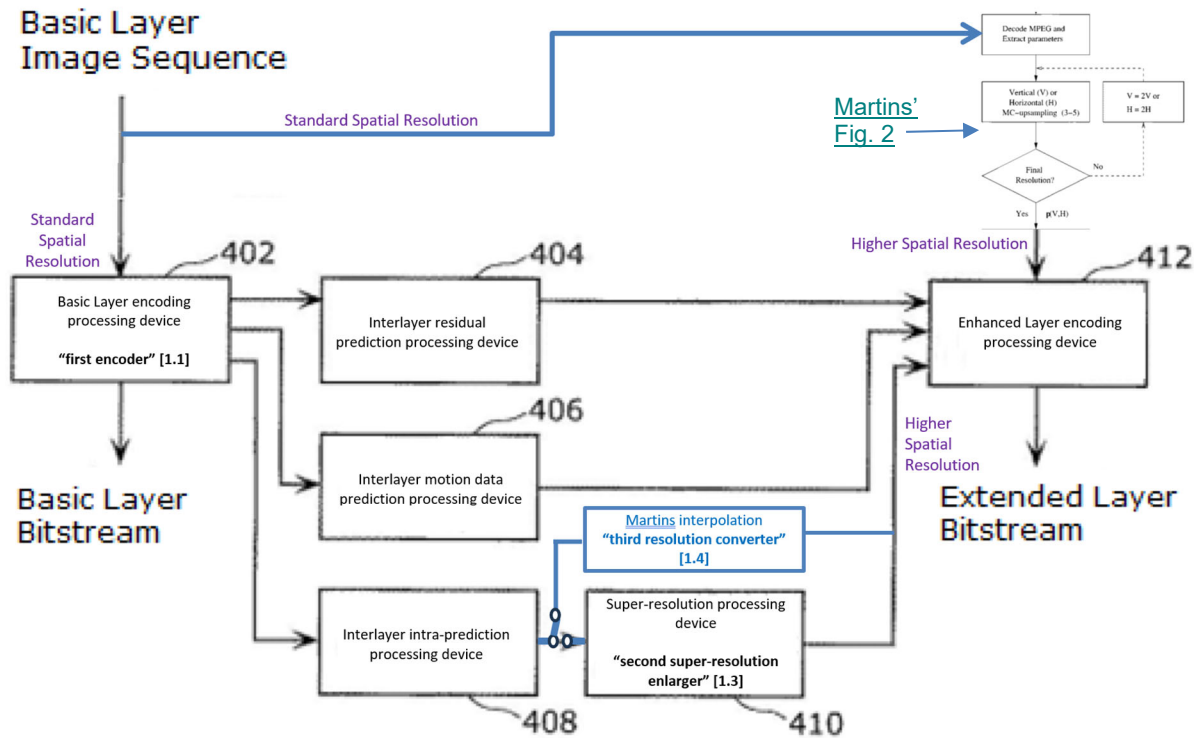


Illustration of using Martins' super-resolution algorithm for generating Extended Layer Image Sequence (second example)

238. For simplicity, Martins' teachings of generating a higher spatial resolution image sequence via super-resolution processing, as taught, e.g., by Martins' Fig. 2, are represented by the illustration below. This simplified illustration is used throughout the limitation-by-limitation analysis that follows.

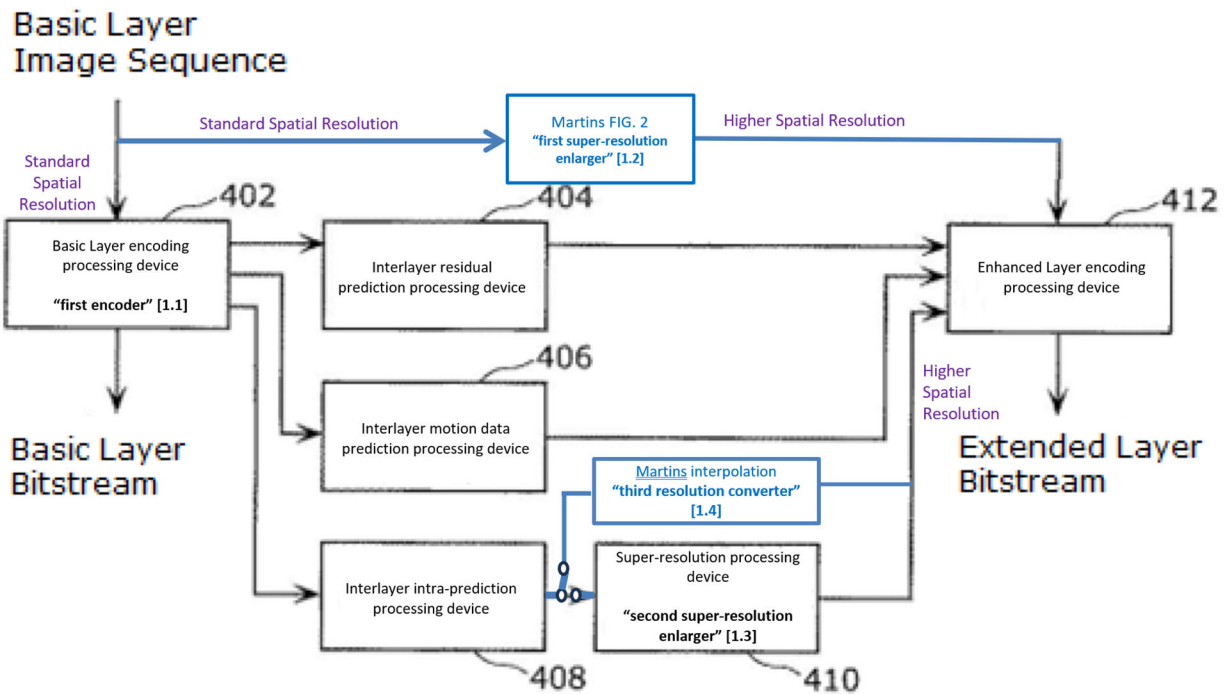


Illustration of Phek-Martins-He Combination (second example).

2. Obvious to Upscale Phek's Reference Pictures and Encoding Target Pictures to the Same Spatial Resolution

239. The challenged claims require that a “spatial resolution” of multiple sets of pictures “are made equal.” For the reasons discussed below, based on the teachings of Phek and Martins, it would have been obvious in the Phek-Martins-He combination to ensure that the pictures provided to the enhanced layer encoding processing device 412 of Phek (which corresponds to the recited “third encoder”) are all the same spatial resolution.

240. Phek discloses a “hierarchical encoding” system “that uses interlayer prediction” between video streams coded at a base layer and an enhancement layer.

Ex.1005 [Phek], [0014]-[0015], FIGs. 4a-4b. Phek states that “[i]nterlayer prediction is used for hierarchical encoding in order to improve encoding efficiency.” Ex. 1005 [Phek], [0014]. Phek does not expressly disclose (nor does it restrict) the type(s) of hierarchical coding that can be implemented by its system, nor does it expressly disclose or restrict the relative spatial resolutions of the base- and enhancement-layer video streams. *Id.* “One example of hierarchical encoding that uses interlayer prediction is ... scalable video encoding.” Ex. 1005 [Phek], [0014]. A POSITA would have understood that a common form of scalable coding is for transmission of base and enhancement layers differentiated by spatial resolution. *See, e.g.*, §§ II.G.5 (discussing scalability in the context of video coding standards), II.G.6 (H.263 reduced resolution update mode), II.G.7 (H.264 Annex Q, paragraph 109).

241. Phek further discloses that the enhancement layer is encoded using reference images from the base layer. Ex.1005 [Phek], [0014]. Specifically, Phek’s encoding path for generating these reference images for the enhancement layer includes “a super-resolution processing device 410 [that] is used to improve [the] reference images.” Ex.1005 [Phek], [0014]. The use of “super-resolution processing ... improve[s] encoding efficiency or picture quality by improving the reference frames used for motion detection, motion compensation prediction, and display.” Ex.1005 [Phek], [0013]. Phek’s super-resolution enhancement involves

processing of a set of lower- resolution images to generate a super-resolution enlarged image of higher quality and higher spatial resolution than the individual lower-resolution images. Ex.1005 [Phek], [0002] (“Super-resolution processing is a technique for generating high-resolution images from a plurality of low-resolution images.”). Subjecting reference pictures to super-resolution processing yields enlarged images whose spatial resolution is greater than the resolution of the decoding targets in the enhancement layer streams. Ex.1005 [Phek], FIG. 4a, [0014].

242. Phek discloses that, at the encoder, “interlayer intra-prediction processing device 408 uses an intra image of the basic layer or a reconstructed image of the basic layer as an additional reference image for encoding the image of the enhanced layer” and “super-resolution processing device 410 is used to improve these additional reference images.” Ex.1005, [0014]. Based on fundamental encoding techniques, it would have been obvious that inter-layer intra-prediction processing device 408 obtains decoded base layer images as reference pictures for the interlayer prediction of enhancement layer images. For example, Phek’s paragraph [0014] states: “interlayer intra-prediction processing device 408 uses an intra image of the basic layer or a **reconstructed image of the basic layer** as an additional reference image for encoding the image of the enhanced layer.” Ex. 1005 [Phek], [0014])

243. Further, as will be discussed below, He teaches that loop filters can be selectively applied to filter/enhance pictures. Ex.1008, [0034]; *supra*, §IV.A.4.b. A POSITA would have recognized that super-resolution processors are another example of a loop filter, and would have recognized that He's teachings could be applied to a super-resolution processing loop filter to selectively control application of super-resolution filtering, such that in an enhancement layer of a layered video coding system, super-resolution filtering would be selectively applied. When super-resolution processing (i.e., super-resolution filtering) is not applied (bypassed, i.e., selectively **not** applied) to a sequence of images, enhanced layer encoding processing device 412 would use the decoded base layer sequence of images (rather than super-resolution processed/filtered versions) acquired by inter-layer intra-prediction processing device 408 as a set of first reference pictures. *Id.*

244. However, because the decoded base layer sequence of images acquired by inter-layer intra-prediction processing device 408 are images at the standard spatial resolution, a POSITA would have recognized that, for use by enhanced layer encoding processing device 412, those images would need to be at the higher spatial resolution. Thus, it would have been obvious to upscale, both the super-resolution ("SR") filtered **and** unfiltered reference pictures that Phek's enhancement layer uses for prediction, to a spatial resolution equal to the spatial resolution of the extended layer image sequence used by the enhanced layer coding processing device 412 as

the encoding target stream (see Technical Background). That is, where it is determined that base layer spatial resolution images from inter-layer intra-prediction processing device 408 should not be super-resolution filtered (enhanced), it would have been obvious to upscale those images to the higher spatial resolution. In particular, a POSITA would have known before the '448 Patent that the reference images used by enhanced layer coding processing device 412 would need to be at the same spatial resolution, whether those reference images were super-resolution enhanced or not. In the latter case, when a reference image is not super-resolution enhanced, resolution conversion to the spatial resolution used by enhanced layer coding processing device 412 would have been obvious, given the well-known nature of resolution conversion at the time of the '448 Patent. A POSITA would have sought to upscale Phek's reference and target images to obtain similar benefits. *See, e.g.,* Ex.1021 [Lan], [0003], [0011], [0016]-[0029], Abstract, FIGS. 2-4; Ex.1020 [Brusewitz], Abstract, [0002], [0026], FIG. 4; Ex.1012 [Park], [0002]-[0003], [0006], [0026]-[0027]).

245. For example, both Phek and Martins themselves include teachings of resolution conversion (e.g., interpolation) of pixels when super-resolution processing does not occur. Phek, for example, teaches that its "super-resolution processing device 410 also performs residual interpolation." Ex.1005 [Phek], [0014].

246. Martins teaches that its super-resolution algorithm creates a super-resolution picture using motion-compensated upsampling, but “[u]sually, some of the pixels” of the super-resolution picture “are undefined because there was no accurate match (of adequate quality) in any of the reference pictures.” Ex.1007 [Martins], p. 807. Thus, to complete the super-resolution picture, those pixels “are assigned values from an interpolated super-resolution picture ... having the same resolution” as the picture generated by motion-compensated upsampling. *Id.* Although Martins refers to interpolation on a pixel basis, a POSITA would have recognized that if **all** the pixels in an image were interpolated, then the interpolation would be on a picture basis. And, a POSITA would have recognized interpolation as a common technique for resolution conversion consistent with the ’448 Patent, and thus Martins motivates a POSITA to incorporate interpolation techniques (e.g., resolution conversion techniques) within a super-resolution processing system.

247. Further, resolution conversion of the pictures in the combination would have been obvious as a predictable application of known techniques to yield predictable results, for which a POSITA would have had a reasonable expectation of success, given the known use of resolution conversion (e.g., upscaling) techniques on images before the ’448 Patent.

3. Obvious Based on He to Selectively Apply Phek’s/Martins’ Super-Resolution (and Upscaling) Loop Filter

248. The challenged claims also require that the recited “third encoder” control “selection of a set of reference pictures and creates a set of data on the selection of the set of reference pictures to identify a selected set of reference pictures during the process for prediction of the third combination.” For the reasons discussed below, based on the teachings of He, it would have been obvious in the Phek-Martins-He combination to control a selection and create a set of data to identify reference pictures used for encoding.

249. Specifically, a POSITA would have found it obvious based on He to implement Phek’s enhancement layer encoder in the combination to selectively encode information reflecting the use of super-resolution processing or upscaling of reference pictures in the enhancement layer of the Phek-Martins-He combination.

250. As previously discussed, the Phek-Martins teachings render obvious techniques for performing super-resolution processing and upscaling of reference images used for inter-layer prediction of images in the enhancement layer. Ex. 1005 [Phek], [0011]-[0015], FIGS. 3a-3b, 4a-4b; Ex.1007 [Martins], 1-2, 6; *supra*, §VI.A.

251. The super-resolution and upscaling (“SR-U”) operations performed on reference pictures in the enhancement layer encoding path form a type of “loop filter,” which is referred to here as a “super-resolution and upscaling loop filter” (or simply an “SR-U filter”). Like other loop filters known in the art, the Phek-Martins-He SR-U loop filter enhances decoded frames before they are used as references for

the prediction of frames in the enhancement layer. Ex.1022 [EE Times – How Video Compression Works], 10-11; Ex.1008 [He], [0006]. A wide range of loop filters were known in the art before the '448 Patent and the SR-U filter in the proposed combination similarly serves to enhance the quality of reference pictures consistent with other known loop filters. *See, e.g.*, Ex.1008 [He], [0006]; Ex.1015 [Kao], [0068]; Ex.1011 [Lei], [0022]-[0023], FIGS. 2-5; Ex.1048 [Doc144], p. 4 (“the filter must be capable of being switched on or off”)); *supra*, §II.F. In addition, the selective application of filters was well-known in the art prior to the '448 Patent.

252. Phek does not explicitly disclose criteria or conditions for the use of super-resolution processing on reference pictures, nor does Martins explain when SR-U filtering should be applied for loop filtering. As discussed above, however, He describes known costs attendant to loop filtering, which may or may not be justified depending on factors such as the degree of improvement in the encoded image quality and compression that would be realized by use of the loop filter, power consumption, and the like. *Supra*, §VI.C. For example, He explains that an in-loop “deblocking filter[] ... consumes a significant amount of power and processing cycles” and “[g]enerally speaking, the entire deblocking filter ... consumes approximately 8% of the computations at the encoder and approximately 35% at the decoder for H.264/AVC.” Ex.1008 [He], [0006]. “Disabling the entire deblocking filter, therefore, provides a significant reduction of power and processing

complexity.” *Id.* To maintain a “desired tradeoff between video quality and complexity reduction,” He discloses control signals that allow a decoder to selectively enable or disable the loop filter. *Id.*, [0069].

253. A POSITA would have found it obvious and would have been motivated to apply He’s suggestion for selective in-loop filtering to the Phek-Martins combination to achieve benefits like those expressly described in He—including to reduce the power consumption and computational complexity of encoding frames in at least the enhancement layer. Ex.1008 [He], [0028], [0031], [0037], [0076], [0069]. In the combination, and based on He’s teaching, an encoder can determine whether or not to apply an in-loop SR-U filter and encode control flags in frame or slice header(s) can to signal to a decoder whether an in-loop SR-U filter was applied. *Id.*

254. Although He describes selective loop filtering in the context of “deblocking” filters, He’s teachings are readily and obviously applicable to other types of loop filters including SR-U filtering. Indeed, the same concerns that motivated He’s use of selective “deblocking”-type loop-filters—including computational complexity and power consumption concerns—are similarly, if not more, applicable to SR-U filtering especially since SR-U filtering involves processing of more images and imposes an even greater computational burden than traditional deblocking filters. *See, e.g.*, Ex.1007 [Martins], 6 (describing need to

“keep the algorithmic complexity down”); Ex.1012 [Park], 13 (describing importance implementing “an efficient [super-resolution] algorithm that reduces the computational cost”).

255. A POSITA also would have known that super-resolution processing is most effective to improve the quality of an image when “some relative scene motions ... exist from frame to frame.” Ex.1011 [Lei], 2. For example, this “relative scene motions” provides for the registration of multiple images and facilitates super-resolution processing. But in the absence of motion, or if the lower-resolution images from which the super-resolution image is constructed are merely “shifted by integer [pixel] units,” then “no new information can be used to reconstruct” a super-resolution image and super-resolution processing will be less effective. Ex.1011 [Lei], 2. Therefore, when super-resolution processing does not improve the reference picture significantly, imposes an unacceptable increase in computational load, and/or when the power consumption required for super-resolution processing is substantial, it would have been obvious to bypass SR-U filtering based on He’s teachings for selective in-loop filtering. For example, consistent with He’s teaching to use an unfiltered reference picture when loop filtering is disabled, it would have been obvious for the enhancement layer encoder in the combination to use a decoded base layer-frame that has not been SR-U filtered for inter-layer prediction of a corresponding frame in the enhancement layer when the SR-U filter is disabled.

Indeed, both Phek (*see* §VI.A) and Martins (*see* §VI.B) disclose using interpolation without super-resolution processing when super-resolution processing fails, suggesting a motivation to combine with He. *See* Ex.1008 [He], [0026] (“the unfiltered reconstructed frame information URF is passed as the reconstructed information RF” when loop filtering is disabled), FIG. 2.

256. The fact that selective loop filtering was standard in video coders before the '448 Patent and that He specifically describes such techniques in a layered coding system like Phek's underscores the obviousness of the combination and that a POSITA would have had a reasonable expectation of success implementing selective SR-U loop filtering in the combination.

B. Relevance of the Phek-Martins-He Combination to the Claims

257. Based on my review, the combined teachings of Phek in view of Martins and He as described in the preceding section would have rendered each of claims 1-3 of the '448 Patent obvious before the Critical Date.

1. Claim 1

[1.P] A moving picture encoding system that makes an encoding of a sequence of moving pictures with a resolution higher than a standard resolution using moving pictures contents which include a sequence of moving pictures with the standard resolution and do not include a sequence of moving pictures with a resolution higher than the standard resolution, the moving picture encoding system comprising:

258. Phek-Martins-He renders obvious [1.P].

259. For example, Phek's FIG. 4a describes an "encoding device (encoder)" in an "[i]nterlayer predictive" coding system. Ex.1005 [Phek], FIG. 4a, [0014]; *see also id.*, [0011]-[0015], FIGs. 3a-3b, 4b. In an encoder, "the image sequence is intra-coded or inter-coded." Ex.1005 [Phek], [0011]. Phek's Figure 4a, as combined with the teachings of Martins and He as detailed above, includes a "*moving picture encoding system*" as recited.

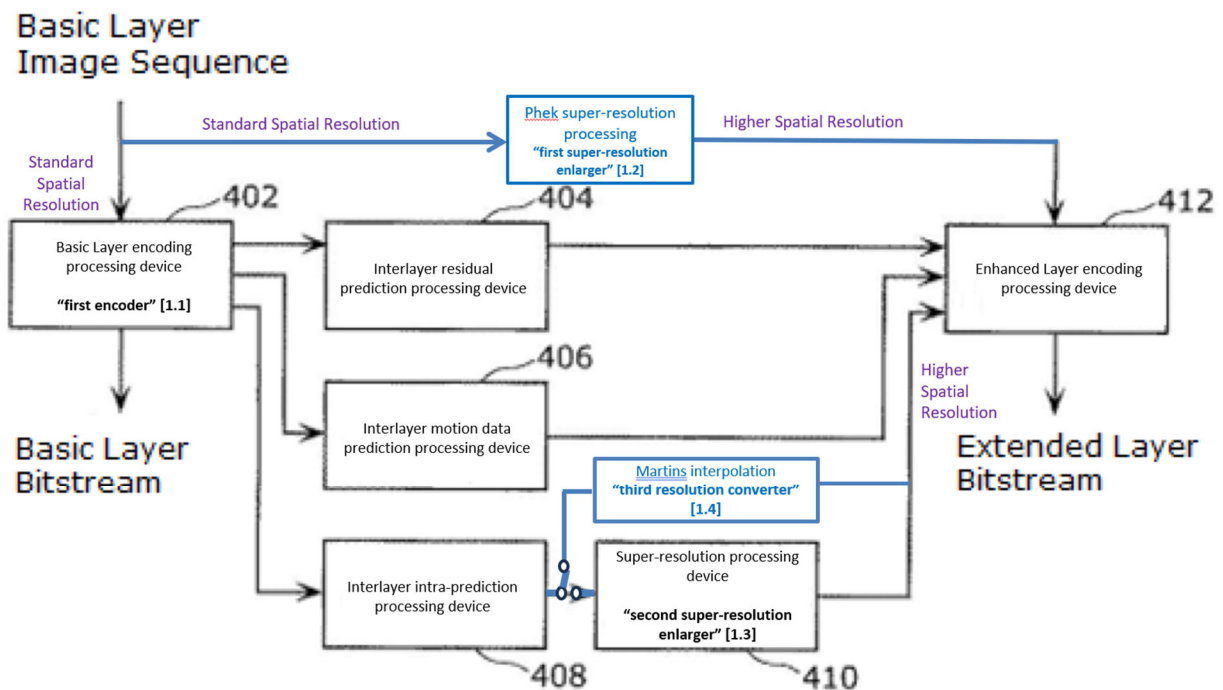


Illustration of Phek-Martins-He Combination (first example).

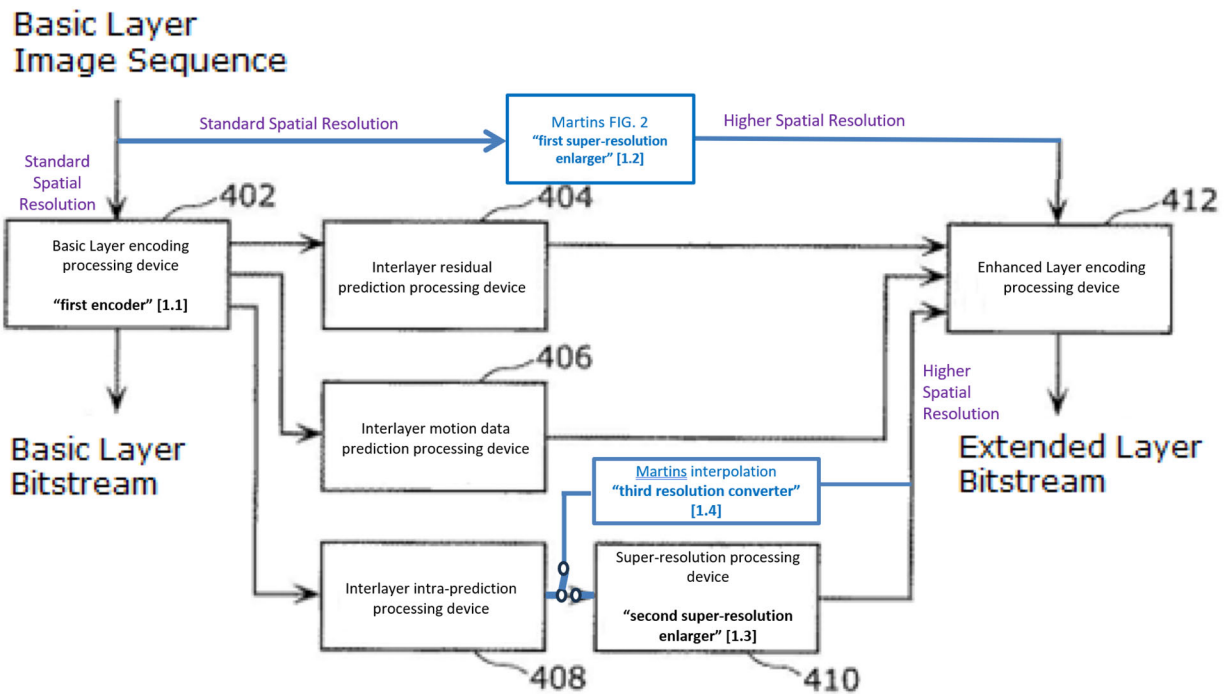


Illustration of Phek-Martins-He Combination (second example).

260. Phek's encoding device in FIG. 4a includes (i) basic layer encoding processing device 402 that processes a basic-layer bitstream to encode a basic layer image sequence, (ii) interlayer intra-prediction processing device 408 that "uses an intra image of the basic layer or a reconstructed image of the basic layer as an additional reference image for encoding the image of the enhanced layer," and (iii) super-resolution processing device 410 that "performs residual interpolation" and "improve[s]" the "additional reference images." Ex.1005 [Phek], [0014].

261. As discussed, Phek does not expressly disclose (nor does it restrict) the type(s) of hierarchical coding that can be implemented by its system. Ex.1005 [Phek], [0014]; *supra*, §VI.A. However, a POSITA understood that a common form

of scalable coding is for transmission of base and enhancement layers differentiated by spatial resolution (“*encoding of a sequence of moving pictures with a resolution higher than a standard resolution*”). For example, Phek’s Figure 4a encodes a basic layer bit stream at a standard spatial resolution and an extended layer bitstream at a higher spatial resolution.

262. Phek’s “basic layer image sequence” corresponds to the recited “*sequence of moving pictures with the standard resolution.*” As shown in the annotated figures above, demonstrating the possible implementation of the combined teachings of Phek, Martins, and He, the extended layer image sequence of Phek may be generated from Phek’s basic layer image sequence (via, e.g., the teachings of either Phek or Martins).

263. Thus, the “*sequence of moving pictures with a resolution higher than a standard resolution*” may be generated “*using moving pictures contents which include a sequence of moving pictures with the standard resolution and do not include a sequence of moving pictures with a resolution higher than the standard resolution,*” that is, only the basic layer image sequence is used, and not any pictures with a higher resolution. *Infra*, §VIII.B ([1.1] and [1.2]).

264. The “*moving picture encoding system*” of the combination as shown in the annotated figures therefore renders obvious the preamble.

[1.1] a first encoder configured to work on a sequence of moving pictures with a standard resolution to implement a first combination of processes for an encoding and a decoding to create a first sequence of encoded bits and a set of decoded pictures with the standard resolution

265. The Phek-Martins-He combination renders obvious [1.1].

266. Specifically, the combination includes Phek's FIG.4a encoder for "hierarchical encoding" which "uses interlayer prediction" where "[a] basic layer image sequence is encoded using a basic layer encoding processing device 402." Ex.1005 [Phek], [0014].

267. Basic layer encoding processing device 402 ("***first encoder***") processes the basic-layer image sequence ("***sequence of moving pictures with a standard resolution***") to implement a process of encoding to create a bitstream for the basic-layer image sequence ("***first sequence of encoded bits***") and decoding to create a decoded set of the basic-layer image sequence for use by interlayer residual prediction processing device 404, interlayer motion data prediction processing device 406, and interlayer intra-prediction processing device 408 ("***set of decoded pictures with the standard resolution***"). See Ex.1005 [Phek], [0014]. The basic layer image sequence is encoded and decoded by basic layer encoding processing device 402 with a first spatial resolution ("***standard resolution***"). Ex.1005, [Phek] [0014] (*see, e.g.*, "reconstructed image of the basic layer"). In the annotated figures

below, I have greyed the portions of Phek's Figure 4a to focus on basic layer encoding processing device 402.

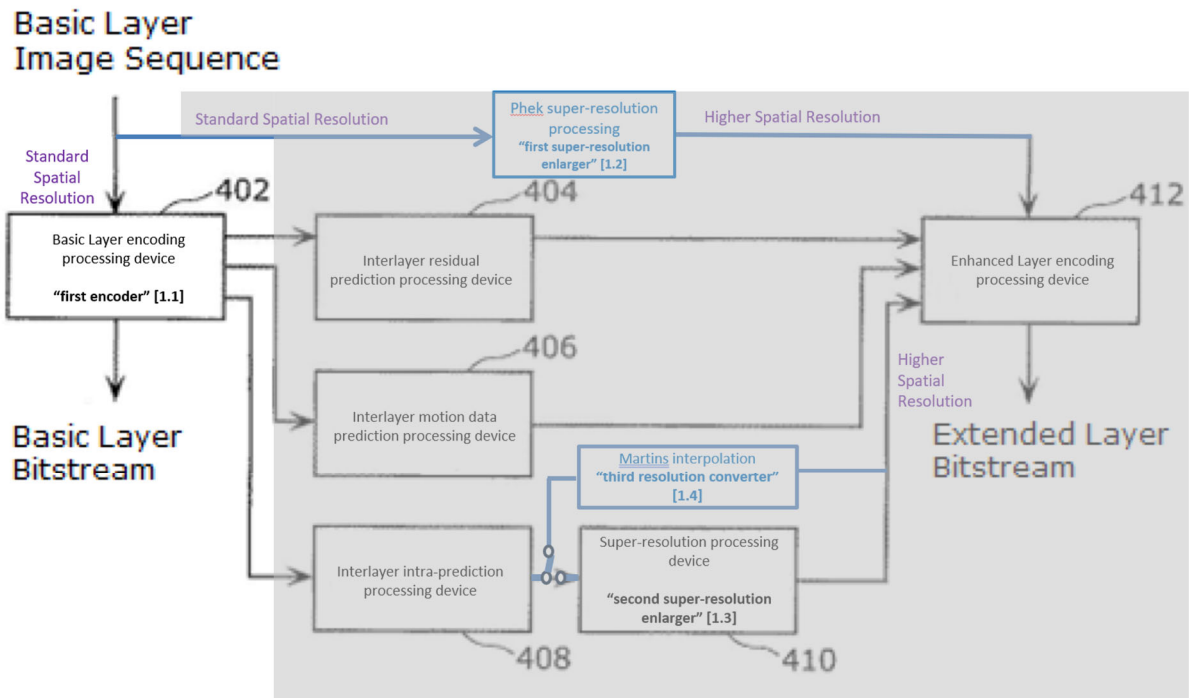


Illustration of Phek-Martins-He Combination (first example).

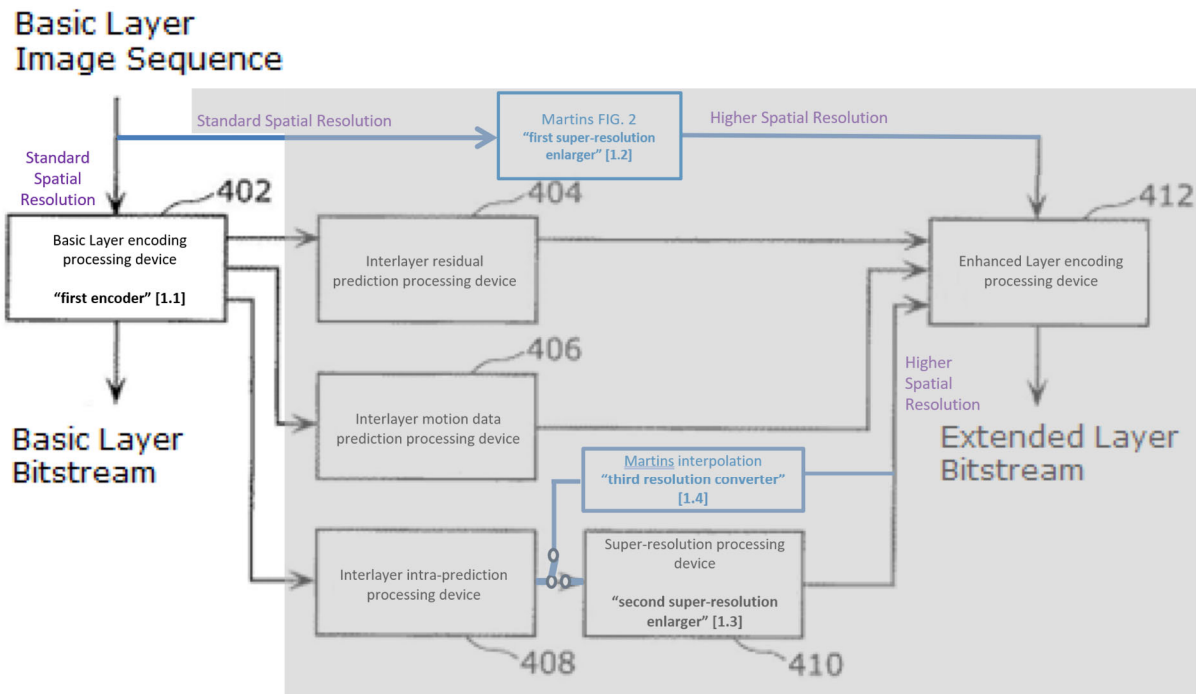


Illustration of Phek-Martins-He Combination (second example).

[1.2] a first super-resolution enlarger configured to work on the sequence of moving pictures with the standard resolution to implement a process for a first super-resolution enlargement to create a set of super-resolution enlarged pictures with a resolution higher than the standard resolution

268. Phek-Martins-He renders obvious [1.2].

269. Phek explains that “[s]uper-resolution processing is a technique for generating high-resolution images from a plurality of low-resolution images.” Ex.1005 [Phek], [0002], *see also* [0003]-[0009]. Phek describes an improved super-resolution processing method with respect to FIGS. 1, 4, and 6. Ex.1005 [Phek], [0016]-[0020]. Applying super-resolution processing to an image enhances the quality and spatial resolution of the image by incorporating into the image additional

information about a scene depicted in other images in the moving picture sequence.
Ex.1005 [Phek], [0011], FIGS. 1, 4, 6a-6b.

270. With respect to Fig. 1a, Phek's "processing devices (functional block in the super-resolution processing device)" include "interpolation processing device 110" which uses interpolation to create "the initial high-resolution estimated image" as well as a "motion detection processing device 102" and "registration processing device 104." Ex.1005 [Phek], [0016], Fig. 1a. The functional blocks of Fig. 1a create a "high-resolution estimated image" and repeatedly process the image "until a predetermined number of cycles has been reached, or until the values of the enhancement coefficients fall below a certain threshold value." Ex.1005 [Phek], [0016], Fig. 1a.

271. Phek's super-resolution technique corresponds to the "enlargement" recited in this limitation: "Super-resolution processing methods include methods that involve image registration. In these methods, motion compensation is performed on observed low-resolution images to convert them into a common high-resolution grid known as a registered image. An initial estimated high-resolution frame is obtained using a spatial interpolation method such as bicubic interpolation." Ex.1005 [Phek], [0003]; *see also supra* §V.B.3.

272. Phek's hierarchical encoding system in Figure 4a includes a super-resolution processing device 410 that operates on decoded pictures from the basic

layer encoder to generate super-resolution enlarged decoded pictures with a resolution higher than the standard resolution. Ex. 1005 [Phek], [0010], [0014], FIG. 4A. While Phek focuses on improving reference images for inter-layer prediction by super-resolution processing, Phek more generally explains that super-resolution processing can also be applied for a variety of purposes including “to improve the resolution and quality of [] inputted image sequences to be displayed” and of “images subject to transcoding and storage.” Ex.1005 [Phek], [0002].

273. Phek’s hierarchical encoding system includes enhanced layer encoding processing device 412 configured to encode an enhanced layer image sequence. Ex.1005 [Phek], [0014], FIG. 4a. Based on Phek’s specific focus on enhancing images through super-resolution processing, and for the reasons discussed in further detail above (*supra* §VIII.A.1), it would have been obvious to a POSITA and consistent with the teachings in Phek for the extended layer image sequence to be formed by applying super-resolution processing to the basic layer image sequence, thereby enhancing the quality and spatial resolution of the extended layer image sequence. Ex.1005 [Phek], [0014], FIG. 4a; §VIII.A.1. Indeed, as detailed above, Phek specifically teaches “use of the super-resolution processing device in a ... pre-processing device” and thus supports using its Figure 1a super-resolution processing device to pre-process an image sequence. Ex. 1005 [Phek], [0010]. And, the ’448 Patent explicitly notes that “the process for the prescribed super-resolution

enlargement may for instance be an application of **any** such super-resolution process” which would include Phek’s Figure 1a super-resolution processing. Ex.1001 [’448 Patent], 22:51-54 (emphasis added).

274. The encoding system would thus employ a super-resolution processing device, e.g., Phek’s super-resolution processing device depicted in Figures 1a or 1b, configured to work on a sequence of moving pictures from the basic layer image sequence with the standard resolution to implement a process for a first super-resolution enlargement to create a set of super-resolution enlarged pictures with a resolution higher than the standard resolution as the extended layer image sequence. Ex.1005 [Phek], [0002]-[0003], [0010], [0016]-[0020], [0014], FIGS. 1-2, 4A, 5, 6A-6B; *supra*, §VIII.A.1.

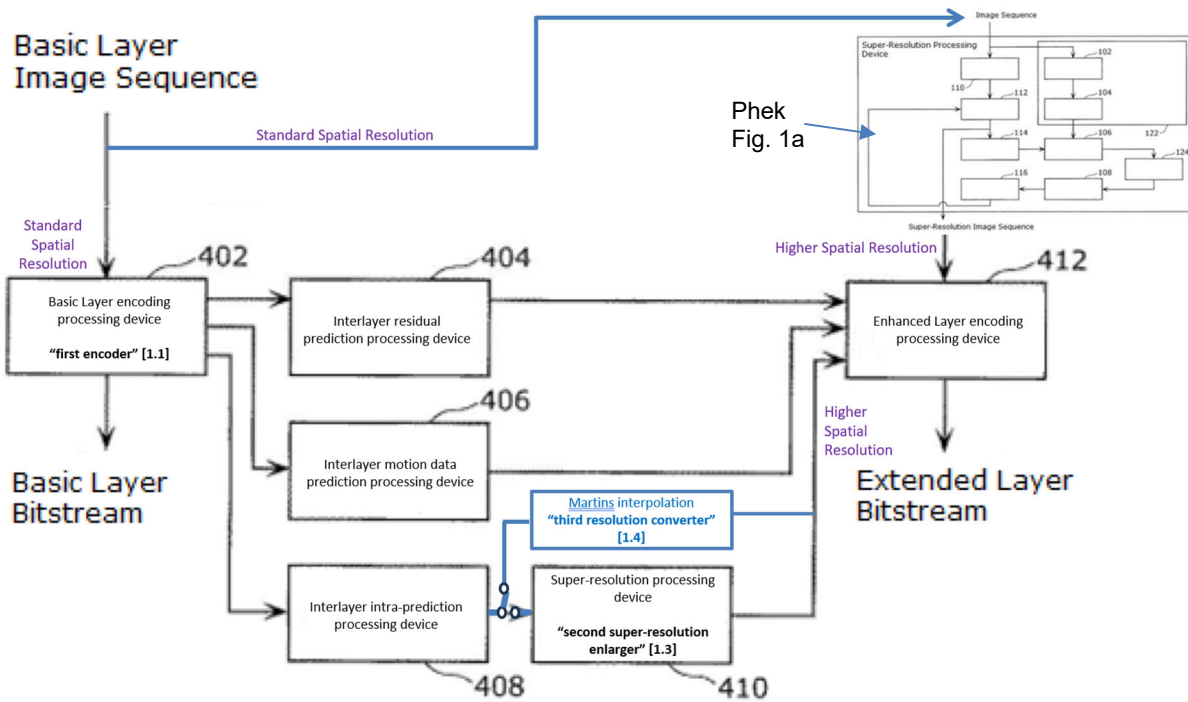


Illustration of Phek’s creation of super-resolution enlarged pictures from the sequence of moving pictures as taught by Fig. 1a (first example) (*see also* simplified version above).

275. Additionally or alternatively, as discussed, Martins’ FIG. 2 discloses “combining multiple pictures to form a single super-resolution picture” using an algorithm “based on motion-compensated spatial upsampling from multiple images.” Ex.1007 [Martins], Abstract; *supra*, §§VI.B; VIII.A.1.

276. Martins explains this algorithm in Section III, “Upsampling to Superresolution Using Motion Compensation,” which includes alignment of pixels of a current picture with pixels of reference pictures using motion estimation, and then combining that information from reference pictures to form a superresolution picture. Ex.1007 [Martins], p. 805. Martins states that the “term *superresolution*

picture is used to refer to the initial MC [motion compensated] upsampled high-resolution image.” Ex.1007 [Martins], p. 805.

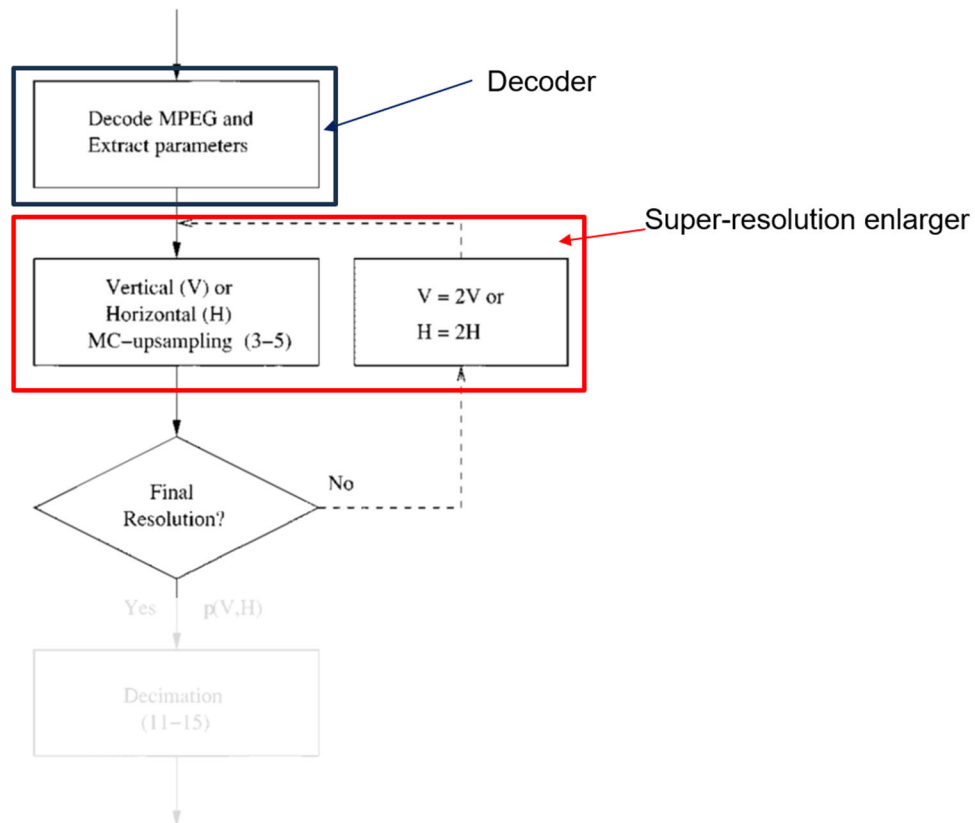


Fig. 2. Overview block diagram. MC upsampling alternates between doubling the resolution vertically and horizontally. Then final step is decimation to the desired format. Equation numbers are given in (). Dashed line marks control flow.

Ex.1007 [Martins], FIG. 2 (amended).

277. In the combination, Martins’ upsampling technique (i.e., enlargement) is applied to the basic layer image sequence (“*sequence of moving pictures with the standard resolution*”) to create the extended layer image sequence in the

enhancement layer “to form a single super-resolution picture” (“*to implement a process for a first super-resolution enlargement to create a set of super-resolution enlarged pictures with a first resolution higher than the standard resolution*”) prior to the encoding processing device 412’s processing. *Id.*; see also Abstract (“obtain a sequence of higher fidelity.”). A POSITA would have understood that Martins’ upsampling technique can be applied to the basic layer image sequence of Phek prior to the encoding processing device 412’s processing, i.e., to generate a sequence at a higher spatial resolution, as shown below.

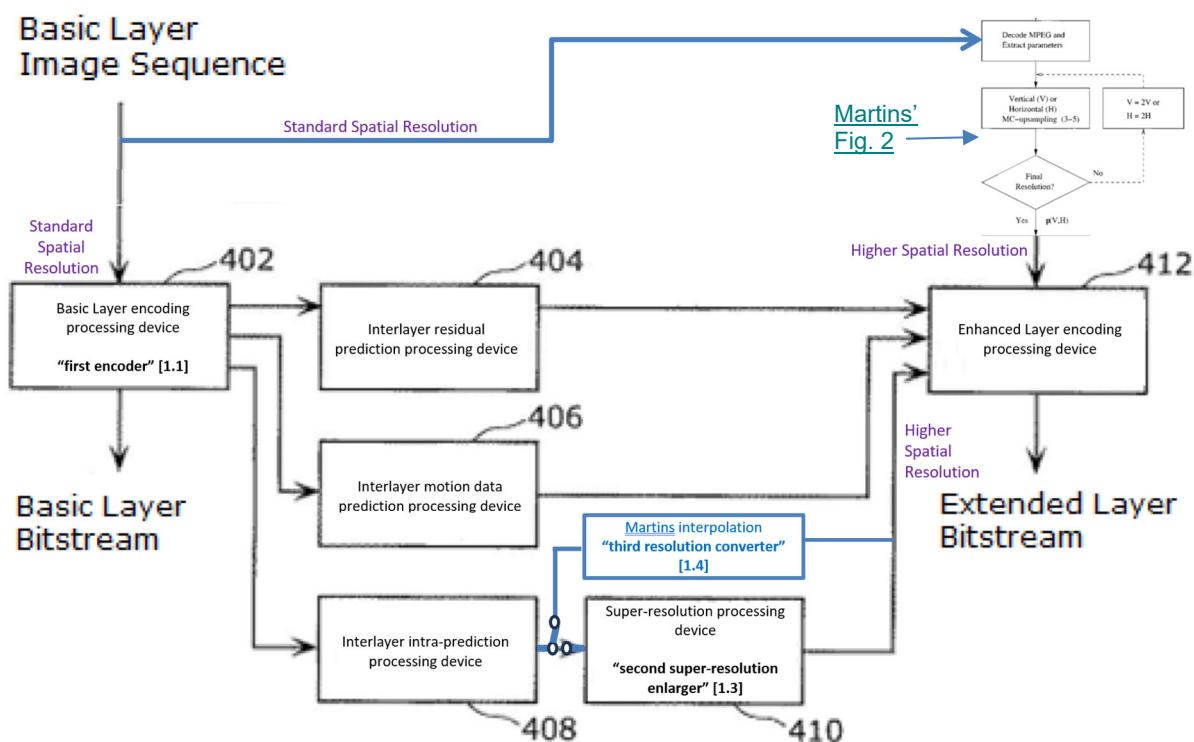


Illustration of using Martins’ super-resolution algorithm for generating Extended Layer Image Sequence (second example) (see also simplified version above).

[1.3] a second super-resolution enlarger configured to acquire the set of decoded pictures from the first encoder to implement thereon a process for a second super-resolution enlargement to create a set of super-resolution enlarged decoded pictures with a resolution higher than the standard resolution

278. Phek-Martins-He renders obvious [1.3].

279. In FIG. 4a, Phek describes a “super-resolution processing device 410 ... used to improve the[] additional reference images” provided by interlayer intra-prediction processing device 408, which “uses an intra image of the basic layer or a reconstructed image of the basic layer for encoding the image of the enhanced layer” based on the basic layer image sequence received by basic layer encoding processing device 402. Ex.1005 [Phek], [0014], FIG. 4a.

280. Phek explains “super-resolution processing...improve[s] encoding efficiency or picture quality by improving the reference frames used for motion detection, motion compensation prediction, and display.” Ex.1005 [Phek], [0013], [0014]. Super-resolution processing device 410 is thus a “***second super-resolution enlarger***” configured to acquire reference images (“***set of decoded pictures with the standard resolution***”) from basic layer encoding processing device 402 (“***first encoder***”), which encoded the basic layer image sequence to produce the set of decoded reference images, through inter-layer intra-prediction processing device to implement an interpolation of pixels with a second enlargement to create “***a set of***

super-resolution enlarged decoded pictures with a second resolution higher than the standard resolution.”

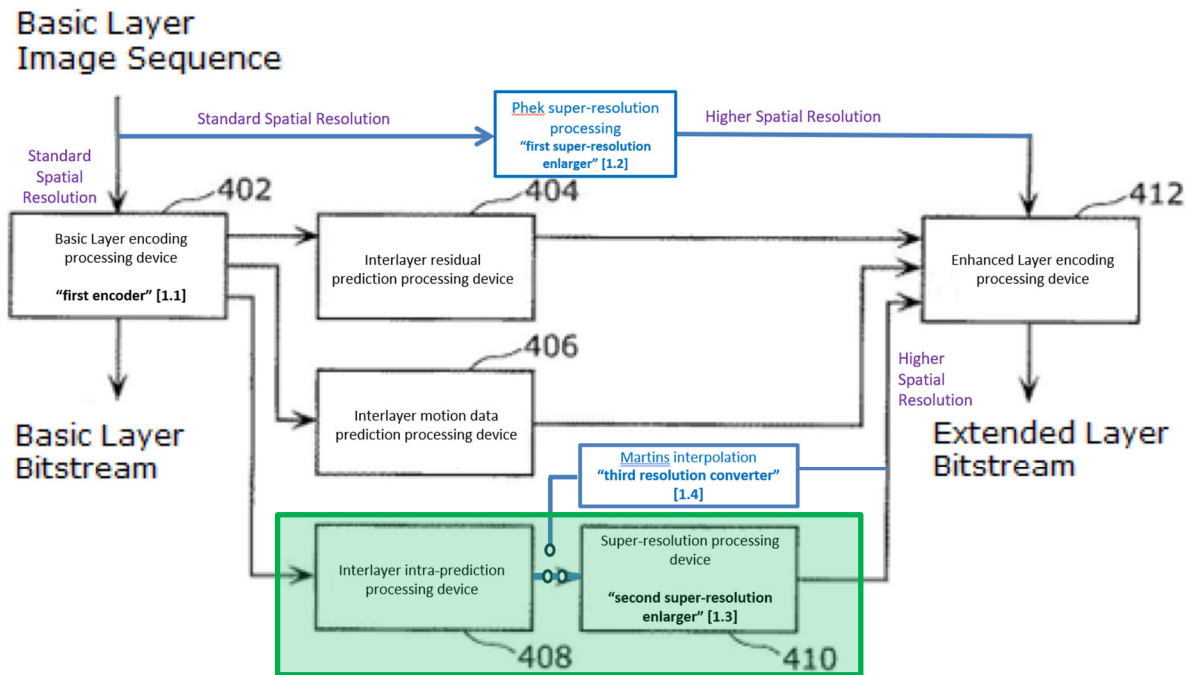


Illustration of Phek-Martins-He Combination (first example).

281. Phek’s super-resolution processing creates a super-resolution enlarged image with improved quality and higher spatial resolution than individual lower-resolution images used as input. *Supra*, §VI.A; Ex.1005, [0002].

282. Thus, Phek’s super-resolution processing device 410 outputs “*a set of super-resolution enlarged decoded pictures with a second resolution higher than the standard resolution.*”

283. Martins’ super-resolution processing and enlargement techniques would also have been obvious to apply in addition to or as an alternative to Phek’s

super-resolution algorithm. *Supra*, §§VI.B, VIII.A.1. Martins' super-resolution enlarger uses "spatial interpolation" to enlarge an image, which improves "quality and spatial resolution of video" by combining multiple images. *Supra*, §§VI.B, VIII.B.1([1.2]); Ex.1007, Abstract, 5, FIG. 2. Martins obtains the super-resolution images using an algorithm "based on motion-compensated upsampling from multiple images and decimation to the desired format." *Supra*, §VI.B; Ex.1007, Abstract.

284. Thus, Martins' Figure 2 algorithm also teaches output of "*a set of super-resolution enlarged decoded pictures with a second resolution higher than the standard resolution.*"

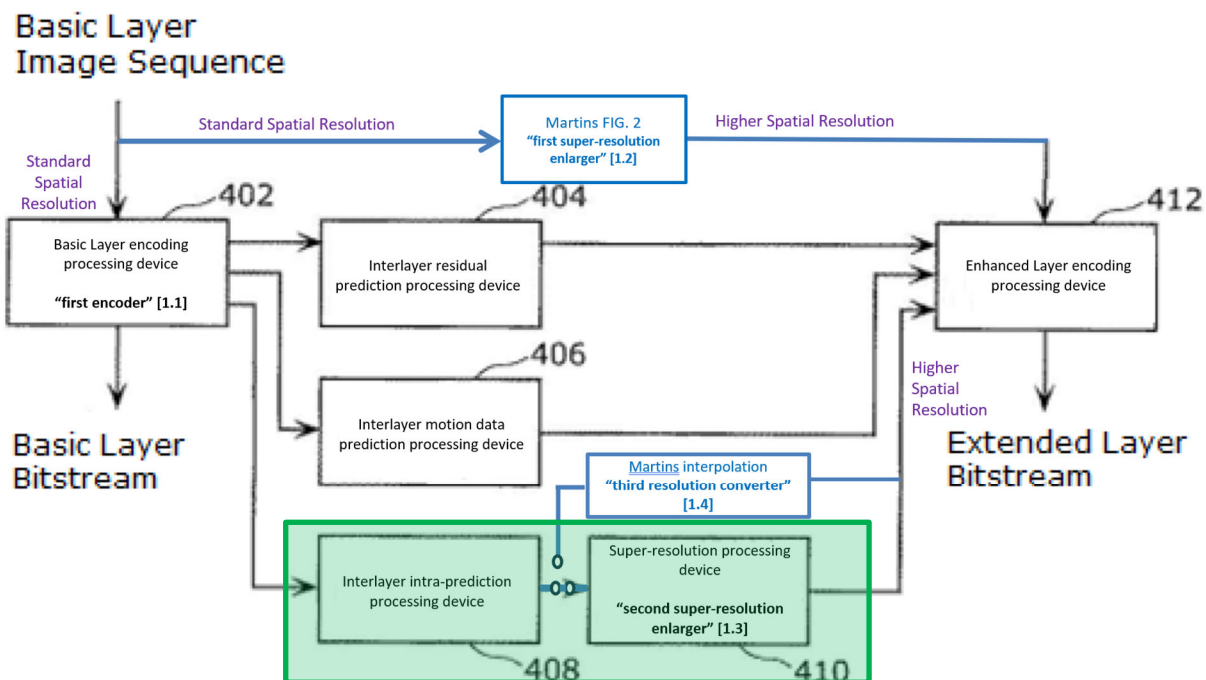


Illustration of Phek-Martins-He Combination (second example).

[1.4] a third resolution converter configured to acquire the set of decoded pictures from the first encoder to implement thereon a process for a third resolution conversion to create a set of resolution converted enlarged decoded pictures with a resolution higher than the standard resolution

285. Phek-Martins-He renders obvious [1.4].

286. Phek discloses that the basic layer encoding processing device 402 (“***first encoder***”) creates a decoded set of the basic-layer image sequence (“***set of decoded pictures with the standard resolution***”). Ex.1005 [Phek], [0014]; *supra*, §[1.1].

287. As discussed, Phek discloses this set of decoded base layer images is used as reference pictures for the interlayer prediction of enhancement layer images. Ex.1005 [Phek], [0014] (“uses ... reconstructed image of the basic layer as an additional reference image for encoding the image of the enhanced layer”); *supra*, §VI.A.

288. In the combined teachings of Phek-Martins-He, based on the teachings of He, a POSITA would have recognized that in some instances, super-resolution processing may not be applied, for example, if super-resolution processing does not improve the reference picture significantly. *See* §VIII.A.2-3. And, as discussed, He teaches that super-resolution and upscaling (“SR-U”) processors, which form a loop filter analogous to He’s deblocking loop filter, can be selectively applied to filter/enhance reference pictures in an enhancement layer of a layered video coding

(e.g., SVC) system. Ex.1008 [He], [0034]. *See also, id.*, [0006]-[0008], [0023]-[0035], [0086]-[0087], FIGs. 1-3, 7; *supra*, §VIII.A.3. Thus, when it is determined that SR-U filtering is not applied (bypassed) to a sequence of images in the proposed combination, Phek's enhanced layer encoding processing device 412 uses the decoded base layer sequence of images (rather than SR-U filtered versions) acquired by inter-layer intra-prediction processing device 408 as a set of reference pictures instead of images processed by super-resolution processing device 410. *See supra*, §VIII.A.3

289. In such cases, a POSITA would have found it obvious that the decoded base layer sequence of images would need to be upsampled (resolution converted) to the same spatial resolution as the encoding target image sequence (i.e., the higher spatial resolution) in order to efficiently serve as reference pictures, and that this upscaling would be done by a “**resolution converter**.” This is consistent with Phek, which discloses the use of interpolation by super-resolution processing device 410 when super-resolution is not used to create missing pixels in the high-resolution picture. *See* Ex.1005 [Phek], [0004-0005], [0016]); *supra*, §VI.A.

290. In addition or alternatively, upscaling, or resolution conversion, of the decoded base layer sequence of images would have been obvious given the teachings of Martins' super-resolution algorithm. For example, like Phek, Martins explains that its algorithm may result in some pixels of the super-resolution picture being

“undefined because there was no accurate match (of adequate quality) in any of the reference pictures.” Ex.1007 [Martins], p. 807. Such pixels are therefore “assigned values from an **interpolated superresolution picture** ... having the same resolution” as the super-resolution picture obtained by motion compensation upsampling. Ex.1007 [Martins], p. 807. A POSITA would have recognized that interpolation, as Martins teaches, is the same method that the '448 Patent itself achieves resolution conversion (*see* Ex.1001, 25:1-41), and thus, Martins’ teachings of assigning values from an interpolated super-resolution picture teaches the “***third resolution converter***” that “***acquire[s] the set of decoded pictures from the first encoder***” (the set of pictures output by basic layer encoding processing device 402) “***to implement thereon a process for a third resolution conversion***” (interpolation) “***to create a set of resolution converted enlarged decoded pictures with a resolution higher than the standard resolution***” (interpolated pictures at the higher spatial resolution) as recited.

Basic Layer Image Sequence

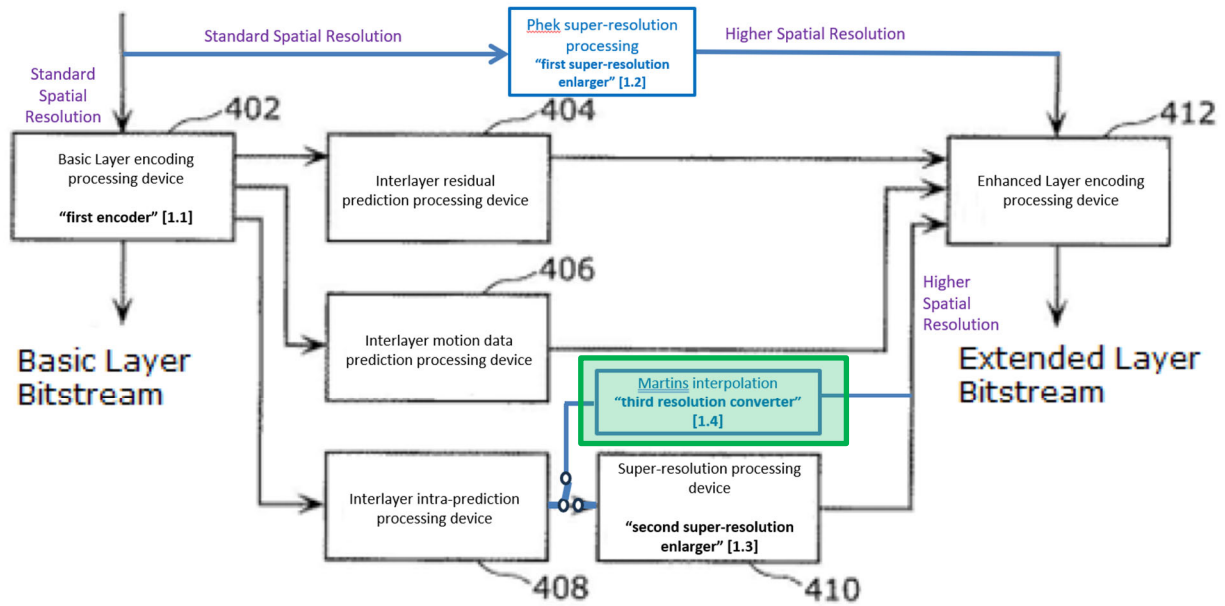


Illustration of Phek-Martins-He Combination (first example).

Basic Layer Image Sequence

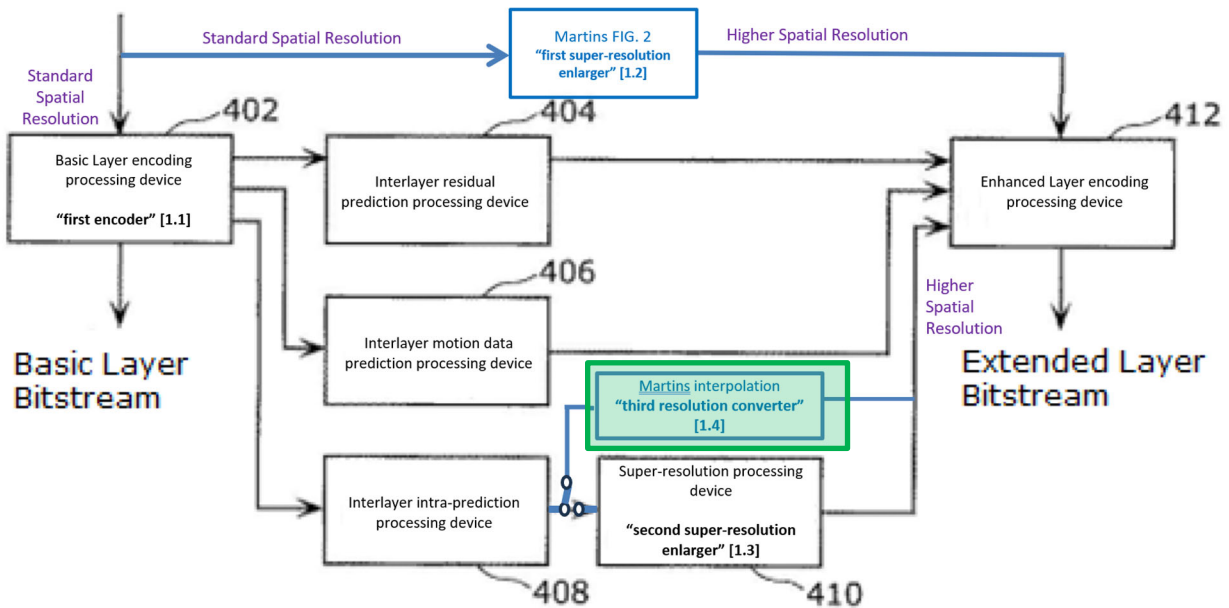


Illustration of Phek-Martins-He Combination (second example).

291. Thus, both proposed examples of the combination render obvious a “*third resolution converter*” as recited.

[1.5] a third encoder configured to have the set of super-resolution enlarged pictures from the first super-resolution enlarger as a set of encoding target pictures, employing the set of super-resolution enlarged decoded pictures from the second super-resolution enlarger and the set of resolution converted enlarged decoded pictures from the third resolution converter as sets of reference pictures, to implement thereon a third combination of processes for a prediction and an encoding to create a third sequence of encoded bits

292. Phek-Martins-He renders obvious [1.5]. As discussed, Phek discloses, and the combination includes, enhanced layer encoding processing device 412 that provides a “*third encoder*.” Ex.1005 [Phek], [0014], FIG. 4a. Enhanced layer encoding processing device 412 encodes the basic layer image sequence received and super-resolution enlarged either by Phek Fig. 1a or Martins Fig. 2 as a set of encoding target pictures, and employing as reference pictures the pictures decoded by basic layer encoding processing device 402, processed through the interlayer intra-prediction processing device 408, and either processed via the super-resolution processing device 410 taught by Phek, the super-resolution device taught by Martins (either of which corresponds to the claimed “*second super-resolution enlarger*”), or the resolution converter as taught by Martins (which corresponds to the claimed “*third resolution converter*”). Ex.1005 [Phek], [0014].

Basic Layer Image Sequence

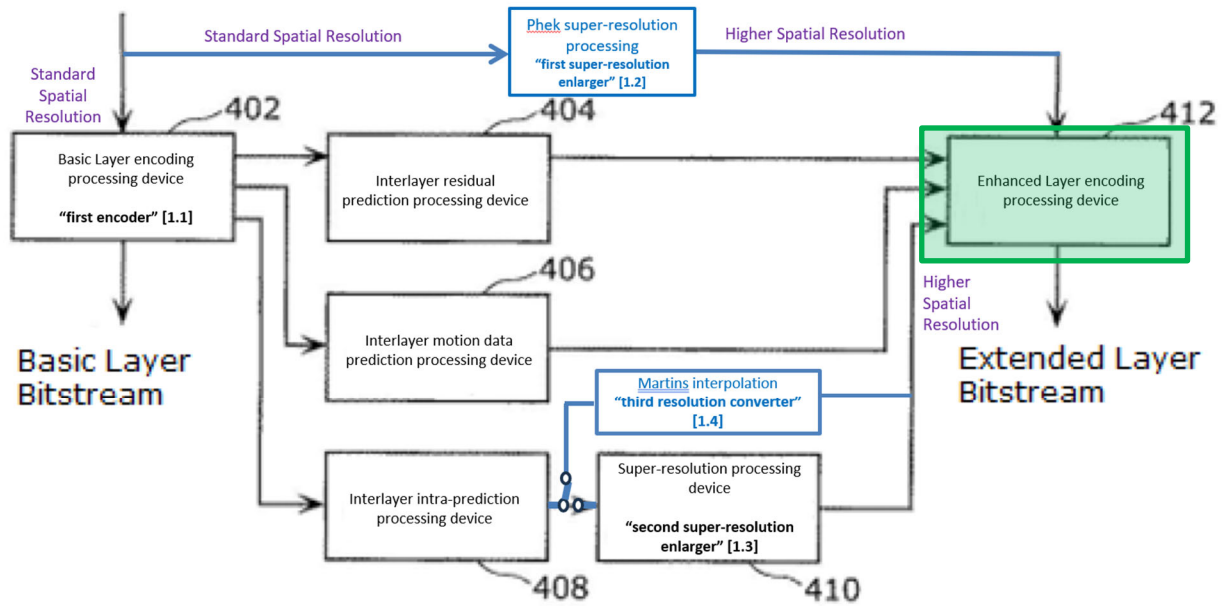


Illustration of Phek-Martins-He Combination (first example).

Basic Layer Image Sequence

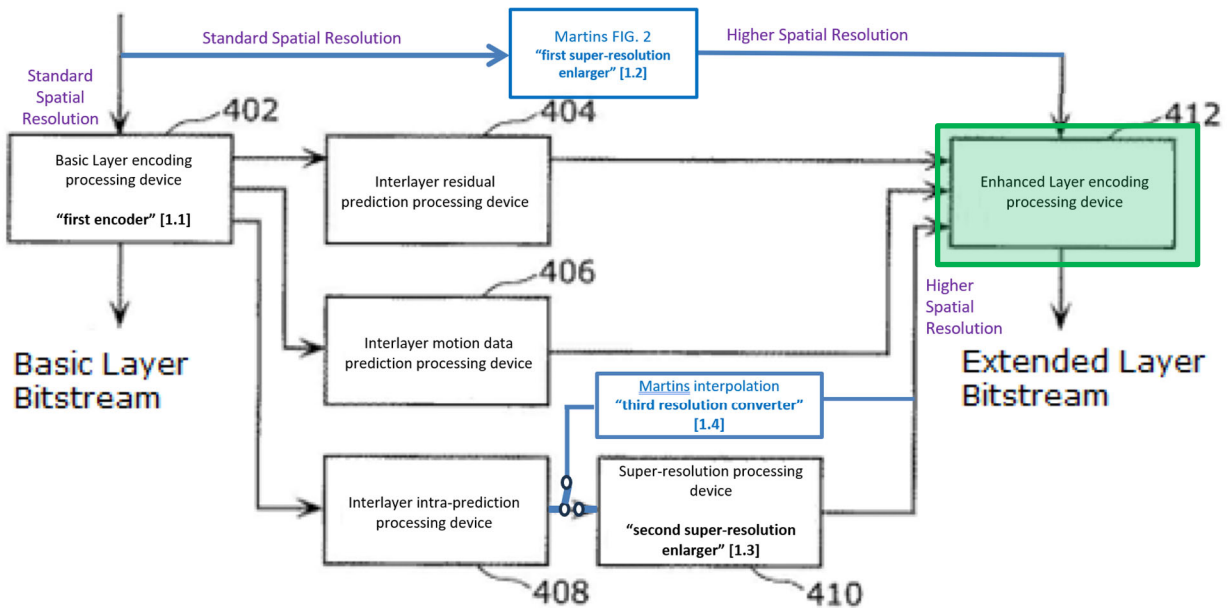


Illustration of Phek-Martins-He Combination (second example).

293. Phek teaches enhanced layer encoding processing device 412 acquires the higher spatial resolution image sequence from the first super-resolution enlarger taught by Phek or Martins (“*set of super-resolution enlarged pictures*”). Ex.1005 [Phek], [0014]. Enhanced layer encoding processing device 412 acquires the higher spatial resolution image sequence as “*a set of encoding target pictures*” allowing the third encoder to encode the enhancement layer stream (e.g., using residual values that represent differences between predicted pictures and target pictures in the enhancement stream). Ex.1005 [Phek], [0014], FIG. 4a. The combination’s third encoder, as shown in the annotated figures, is also configured to acquire (i) “*the set of super-resolution enlarged decoded pictures*” from the second super-resolution enlarger (e.g., super-resolution processing device 410), and (ii) “*the set of resolution converted enlarged decoded pictures*” from the third resolution converter. Ex.1005 [Phek], [0014], FIG. 4a.

294. Phek discloses that, at the encoder, “interlayer intra-prediction processing device 408 uses an intra image of the basic layer or a reconstructed image of the basic layer as an additional reference image for encoding the image of the enhanced layer” and “super-resolution processing device 410 is used to improve these additional reference images.” Ex.1005 [Phek], [0014]. Phek therefore discloses, consistent with fundamental decoding techniques, that inter-layer intra-prediction processing device 408 obtains decoded base layer images as reference

pictures for the interlayer prediction of enhancement layer images. *See also* Ex.1005 [Phek], Fig. 4a.

295. As detailed above in the motivation to combine section, and further in the analysis of limitation [1.6]-[1.7] below, the “*third encoder*” uses “*the set of super-resolution enlarged decoded pictures from the second super-resolution enlarger*” as well as “*the set of resolution converted enlarged decoded pictures from the third resolution converter as sets of reference pictures*” and selects between the two sets for encoding a particular sequence. Selection between the two sets is taught, e.g., by He, which teaches that super-resolution and upscaling processors, which form a loop filter that is analogous to He’s deblocking loop filter, can be selectively applied to filter/enhance reference pictures in an enhancement layer of a layered video coding (e.g., SVC) system. Ex.1008 [He], [0034]. *See also, id.*, [0006]-[0008], [0023]-[0035], [0086]-[0087], FIGs. 1-3, 7; *supra*, §VIII.A.3.

296. When applying SR-U filtering to a sequence of images, enhanced layer encoding processing device 412 acquires SR-U filtered versions of the decoded base layer sequence of images (“*set of super-resolution enlarged decoded pictures from the second super-resolution enlarger*”) as a set of first reference pictures. Ex.1005 [Phek], [0014]-[0015], FIG. 4b; *supra*, §VIII.A.3. When SR-U filtering is not applied (bypassed) to a sequence of images, enhanced layer encoding processing device 412 uses the decoded base layer sequence of images (rather than SR-U

filtered versions) acquired by inter-layer intra-prediction processing device 408 and resolution converted as a set of second reference pictures (*“set of resolution converted enlarged decoded pictures from the third resolution converter”*). *Id.*

297. The combination’s *“third encoder”* thus acquires the SR-U filtered first set of reference pictures when SR-U filtering is enabled and acquires the decoded, resolution converted base layer sequence of images as a set of second reference pictures both when SR-U filtering is enabled (as needed to perform the SR-U filtering) and when it is not.

298. The following two figures illustrate using the SR-U filtered version of decoded base layer images as a set of first reference pictures when the SR-U filter is enabled in the two Phek-Martins-He examples:

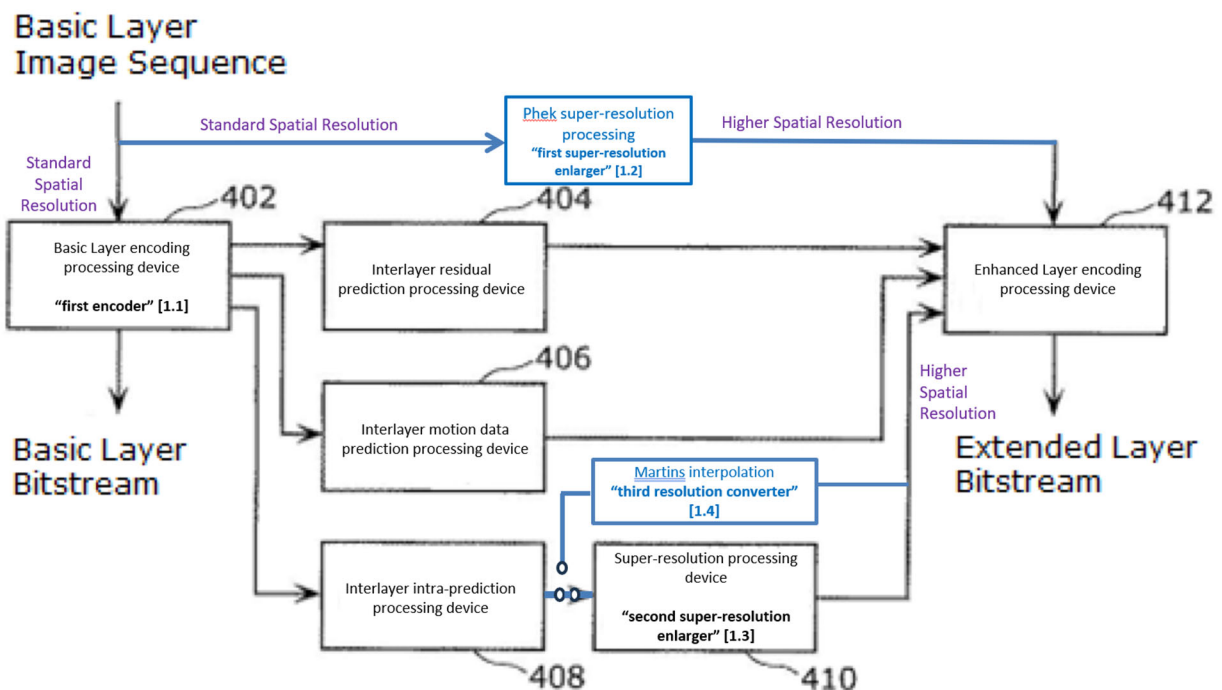


Illustration of Phek-Martins-He Combination (first example) (1st reference picture).

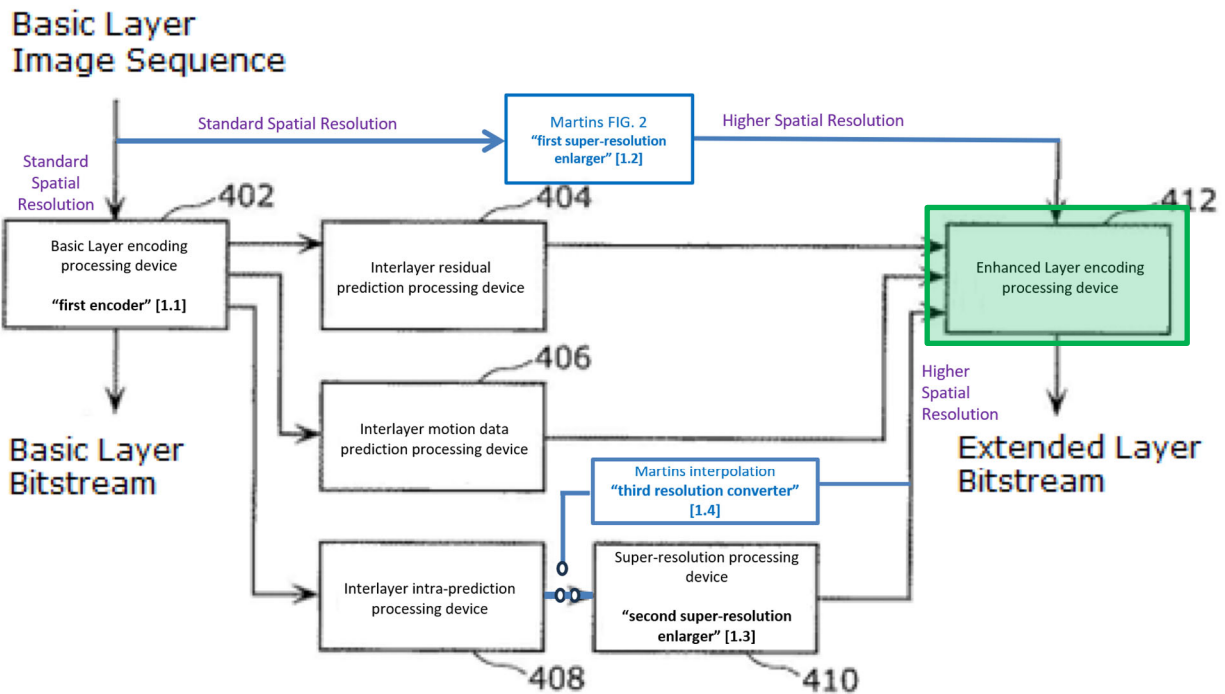


Illustration of Phek-Martins-He Combination (second example) (1st reference picture).

299. By contrast, the following figures illustrate using the non-SR-U filtered version of decoded base layer images as a set of second reference pictures when the SR-U filter is disabled in the two Phek-Martins-He examples:

Basic Layer Image Sequence

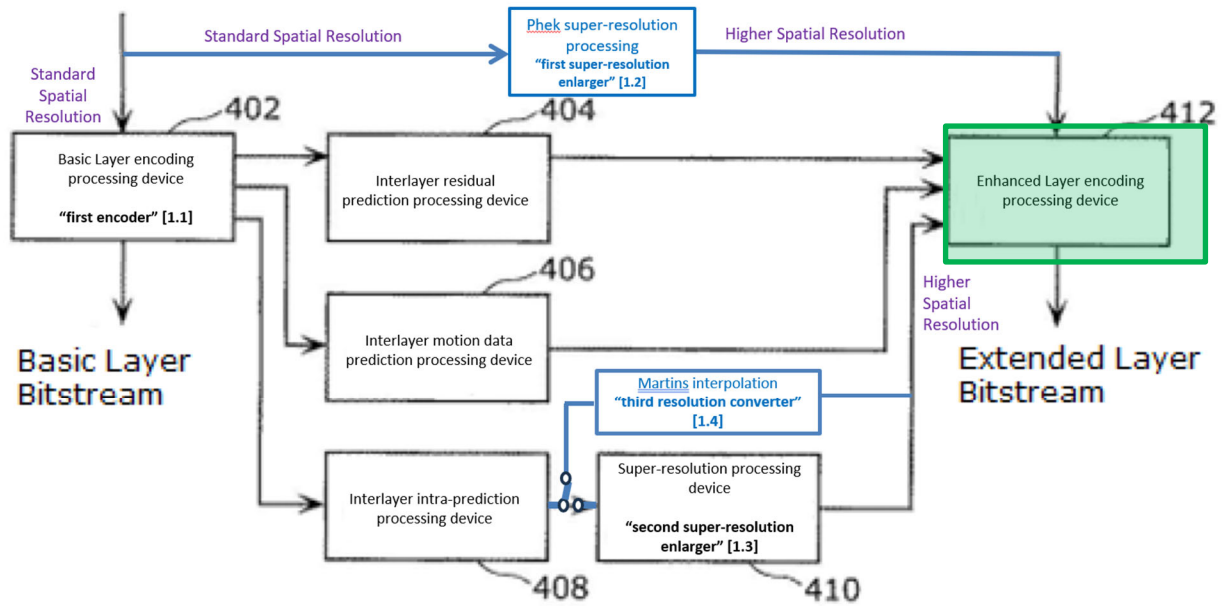


Illustration of Phek-Martins-He Combination (first example) (2nd reference picture).

Basic Layer Image Sequence

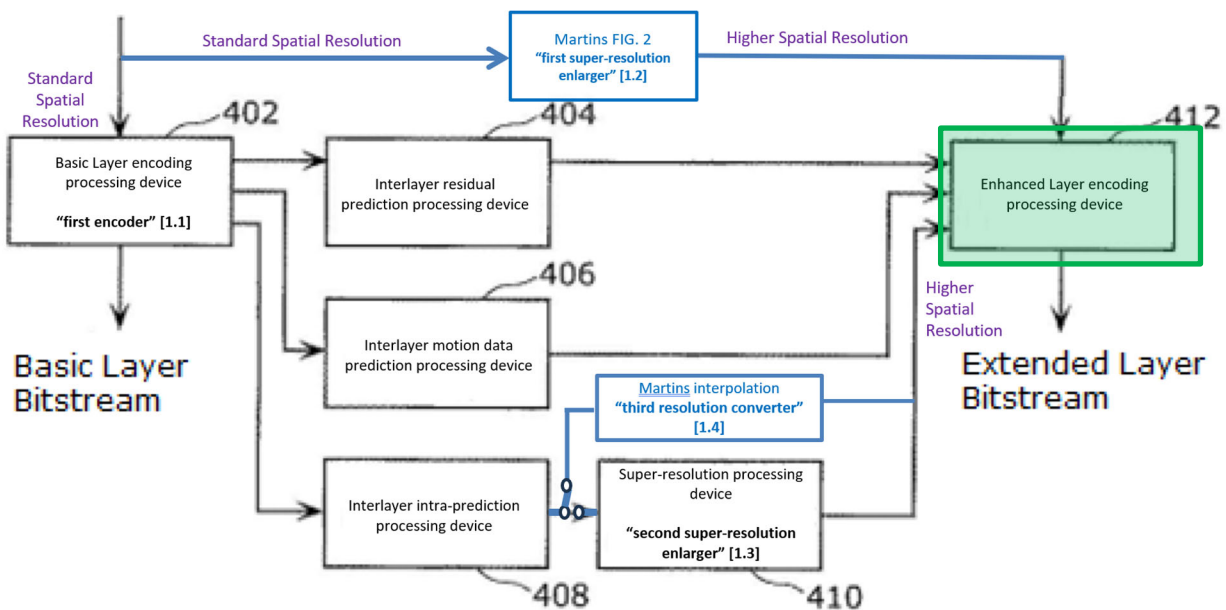


Illustration of Phek-Martins-He Combination (second example) (2nd reference picture).

[1.6] wherein a spatial resolution of the set of super-resolution enlarged pictures, that of the set of super-resolution enlarged decoded pictures, and that of the set of resolution converted enlarged decoded pictures are made equal

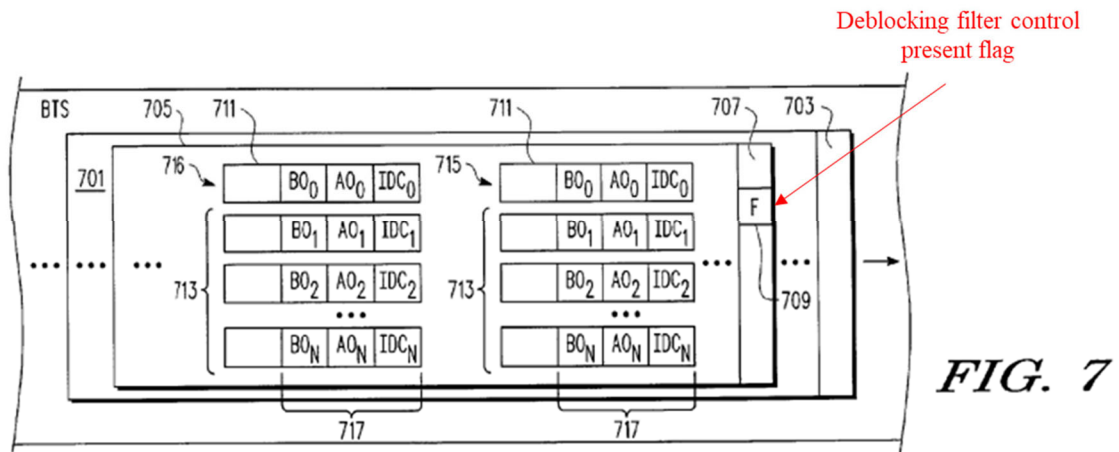
300. Phek-Martins-He renders obvious [1.6]. It would have been obvious to implement the Phek-Martins-He combination's encoding device such that the set of encoding targets, set of first reference pictures (e.g., decoded base layer pictures with SR-U filtering), and set of second reference pictures (e.g., resolution converted enlarged decoded base layer pictures without SR-U filtering) have the same resolution. *Supra*, §§VIII.A.2, VI.B. For example, as detailed in the motivation to combine section above, doing so would have been consistent with the well-known operation of encoding devices prior to the '448 Patent.

[1.7] wherein the third encoder controls a selection of a set of reference pictures and creates a set of data on the selection of the set of reference pictures to identify a selected set of reference pictures during the process for prediction of the third combination

301. Phek-Martins-He renders obvious [1.7]. The combination's enhancement layer encoding device 412 encodes pictures in the enhancement layer bitstream using either decoded base layer pictures enhanced by SR-U filtering ("***the set of super-resolution enlarged pictures from the first super-resolution enlarger***") or decoded base layer pictures not enhanced by SR-U filtering ("***the set of resolution converted enlarged decoded pictures from the third resolution converter***").

302. Enhancement layer encoding processing device 412 (and/or other components in the enhancement layer corresponding to the “*third encoder*”) selects whether to use SR-U filtered or non-filtered reference pictures based on control information, such as a MINSAD (minimum sum of absolute differences) value used to measure complexity of the video content, motion vectors, or a power saving factor (PSF), based on He. Ex.1008 [He], [0029]-[0031], FIG. 1; *see also supra*, §VIII.A.3.

303. He discloses “a deblocking control flag of scalable video system for indicating presentation of deblocking parameters for multiple layers” in a bitstream. Ex.1008 [He], [0002]. “The deblocking parameters determine whether deblock filtering is enabled or disabled” where the parameter is “set to ‘0’ to enable deblock filtering and to ‘1’ to disable deblock filtering.” Ex.1008, [0007]. He’s FIG. 7 illustrates “deblocking filter control information provided in an output bitstream BTS ... or provided in an input bitstream BTS (e.g., IBTS) to the video decoder 603 ... The bitstream BTS includes at least one video information sequence 701 ... The video information sequence 701 further includes one or more video frames including an exemplary video frame 705 ... The video frame 705, for example, includes a frame header 707 at the front of the video frame 705 which further includes a deblocking filter control present flag 709, labeled ‘F’.” Ex.1008 [He], [0052].



Ex.1008, FIG. 7

304. He's FIG. 2 discloses deblock control circuit 221 which, based on the deblocking control flag, issues control signals to deblocking filter 131 to selectively enable/disable use of a deblock filter on reconstructed (decoded) frames. Ex.1008, FIG. 2, [0006]-[0008]. When the deblock filter is enabled, the filtered (enhanced) reconstructed frame is used as a reference picture for prediction of a subsequent frame; when disabled, the original (unfiltered or non-enhanced) reconstructed frame is used as the reference picture for prediction of a subsequent frame. *Id.*, [0034].

305. He's "deblocking filter 211 is configured in a similar manner as the deblocking filter 131, and includes a similar decoder filter control DFC signal provided by a deblock control circuit 221....if the deblocking filter 211 is disabled, then the unfiltered reconstructed frame information URF' is passed as the "filtered" reconstructed information RF' in which deblock filtering is bypassed." Ex.1008 [He], [0034].

306. As detailed above, He further explains that “the video encoder 100 provides control information within the bitstream BTS to cause the deblock control circuit 221 of the video decoder 200 to provide the DFC signal to adaptively disable deblock filtering for a corresponding slice or frame of URF’ information being processed in the video decoder 200.” Ex.1008 [He], [0034], [0008 (enable/disable parameter)]. Thus, He teaches “*data on the selection*” as recited.

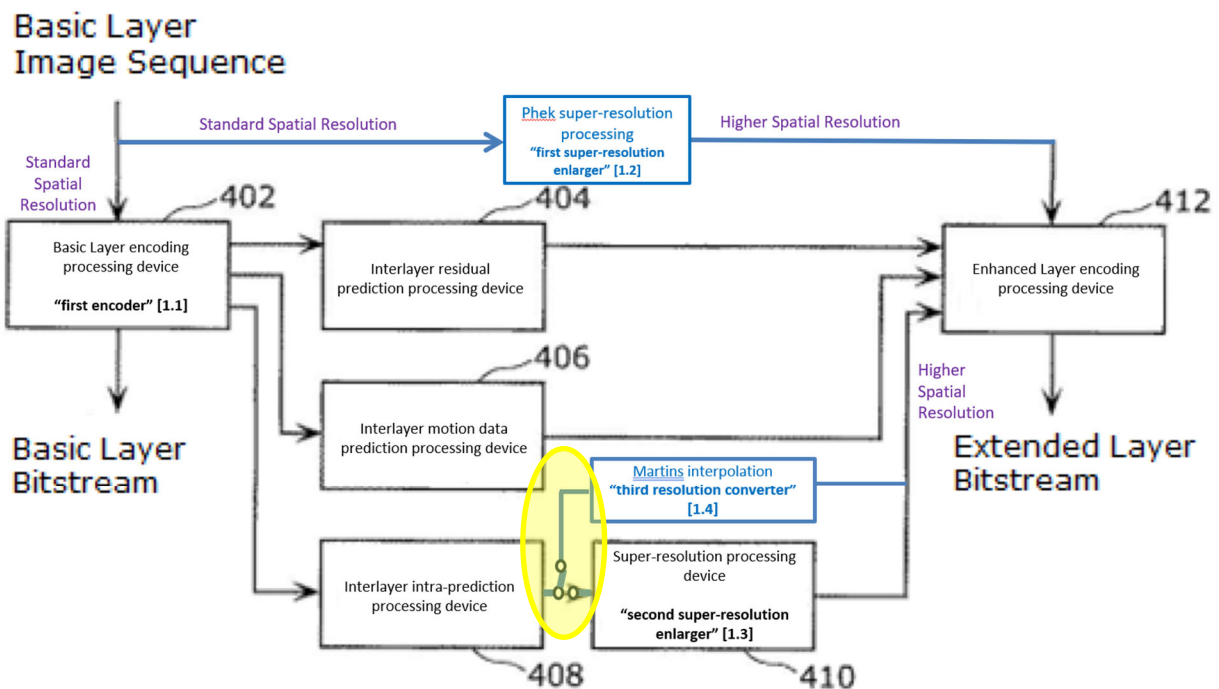


Illustration of Phek-Martins-He Combination (first example) (reference picture selection).

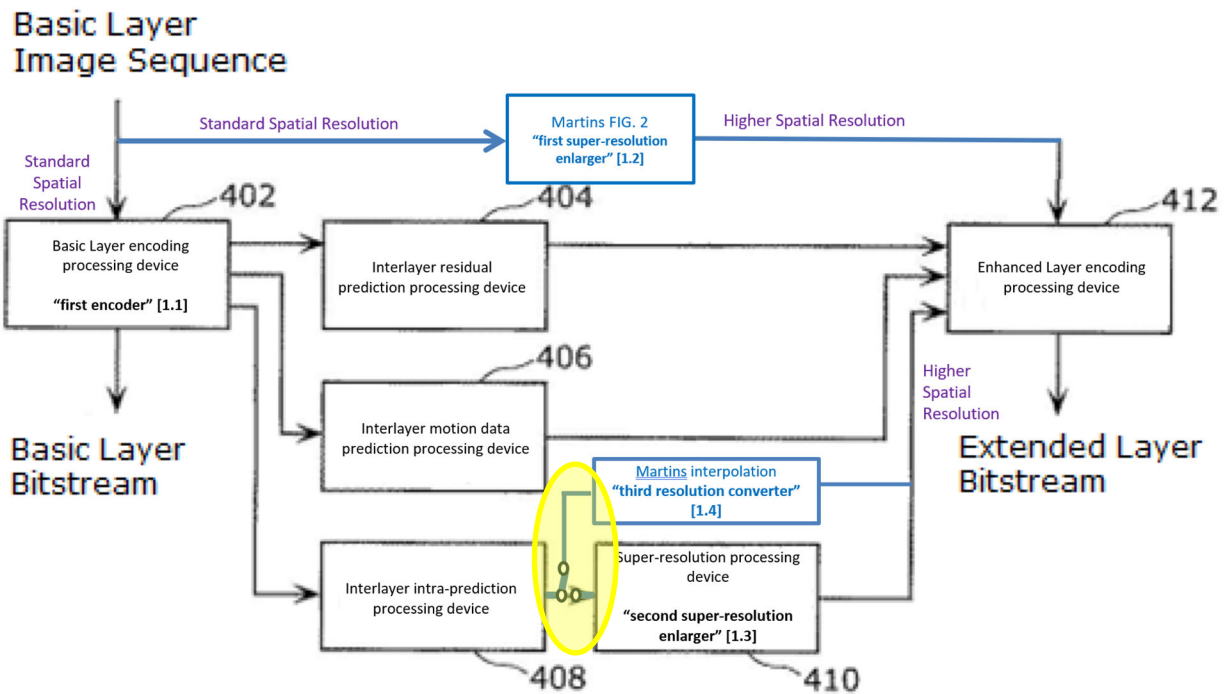


Illustration of Phek-Martins-He Combination (second example) (reference picture selection).

307. These figures illustrate selectivity of the SR-U filter in yellow.

308. As illustrated above, the **"third encoder"** as implemented in the Phek-Martins-He combination selects either the super-resolution enlarged base-layer decoded pictures sequence (when SR-U filter is enabled) as the set of first reference pictures or the resolution converted enlarged base-layer decoded pictures sequence (when SR-U filter is disabled) as a set of second reference pictures. Control signals like He teaches (e.g., analogous to He's "deblocking enable/disable parameter" and "deblocking control flag") would thus be created by the **"third encoder"** in the combination to provide an enhancement-layer decoding path with **"a set of data on**

the selection of the set of reference pictures to identify a selected set of reference pictures.”

309. This is consistent with Martins’ teachings that “[e]ach pixel of the resulting superresolution picture ... is assigned the attribute of whether it was determined by motion compensation” (i.e., whether the pixel belongs to a super-resolution picture obtained by motion-compensated super-resolution processing) “... or interpolation” (i.e., whether the pixel belongs to an interpolated super-resolution picture). Ex.1007 [Martins], p. 807. In other words, Martins teaches that super-resolution processing is enabled or disabled on a pixel-by-pixel basis, and an attribute (i.e., “*data*”) is assigned to determine whether a pixel is interpolated or super-resolution processed, as in the proposed combination on a picture basis.

2. Claim 2

[2.P] A moving picture encoding system that makes an encoding of a sequence of moving pictures with a resolution higher than a standard resolution using moving pictures contents which include a sequence of moving pictures with the standard resolution and do not include a sequence of moving pictures with a resolution higher than the standard resolution, the moving picture encoding system comprising:

310. The Phek-Martins-He combination renders obvious this claim element for at least the reasons described above with respect to [1.P].

[2.1] a first encoder configured to work on a sequence of moving pictures with a standard resolution to implement a first combination of processes for an encoding and a decoding to create a first sequence of encoded bits and a set of decoded pictures with the standard resolution

311. The Phek-Martins-He combination renders obvious this claim element for at least the reasons described above with respect to [1.1].

[2.2] a first super-resolution enlarger configured to work on the sequence of moving pictures with the standard resolution to implement a process for a first super-resolution enlargement to create a set of super-resolution enlarged pictures with a resolution higher than the standard resolution

312. The Phek-Martins-He combination renders obvious this claim element for at least the reasons described above with respect to [1.2].

[2.3] a third resolution converter configured to acquire the set of decoded pictures from the first encoder to implement thereon a process for a third resolution conversion to create a set of resolution converted enlarged decoded pictures with a resolution higher than the standard resolution

313. The Phek-Martins-He combination renders obvious this claim element for at least the reasons described above with respect to [1.3].

[2.4] a third encoder configured to have the set of super-resolution enlarged pictures from the first super-resolution enlarger as a set of encoding target pictures, employing the set of resolution converted enlarged decoded pictures from the third resolution converter as a set of reference pictures, to implement thereon a third combination of processes for a prediction and an encoding to create a third sequence of encoded bits

314. The only difference between [2.4] and [1.5] is that [2.4] only requires one set of reference pictures, namely “the set of resolution converted enlarged decoded pictures from the third resolution converter.” Thus, the combination renders obvious this limitation for the same reasons as described for [1.5]. *Supra* [1.5].

[2.5] wherein a spatial resolution of the set of super-resolution enlarged pictures, and that of the set of resolution converted enlarged decoded pictures are made equal

315. The only difference between [2.5] and [1.6] is that [2.5] only requires one set of reference pictures, namely “the set of resolution converted enlarged decoded pictures” to be made equal to the spatial resolution of the “the set of super-resolution enlarged pictures.” Thus, the combination renders obvious this limitation for the same reasons as described for [1.6]. *Supra* [1.6].

[2.6] wherein the third encoder controls a selection of a set of reference pictures and creates a set of data on the selection of the set of reference pictures to identify a selected set of reference pictures during the process for prediction of the third combination

316. The Phek-Martins-He combination renders obvious this claim element for at least the reasons described above with respect to [1.7].

3. Claim 3

[3.P] A moving picture encoding system that makes an encoding of a sequence of moving pictures with a resolution higher than a standard resolution using moving pictures contents which include a sequence of moving pictures with the standard resolution and do not include a sequence of moving pictures with a resolution higher than the standard resolution, the moving picture encoding system comprising

317. The Phek-Martins-He combination renders obvious this claim element for at least the reasons described above with respect to [1.P].

[3.1] a first encoder configured to work on a sequence of moving pictures with a standard resolution to implement a first combination of processes for an encoding and a decoding to create a first sequence of encoded bits and a set of decoded pictures with the standard resolution

318. The Phek-Martins-He combination renders obvious this claim element for at least the reasons described above with respect to [2.1].

[3.2] a first super-resolution enlarger configured to work on the sequence of moving pictures with the standard resolution to implement a process for a first super-resolution enlargement to create a set of super-resolution enlarged pictures with a resolution higher than the standard resolution

319. The Phek-Martins-He combination renders obvious this claim element for at least the reasons described above with respect to [2.2].

[3.3] a second super-resolution enlarger configured to acquire the set of decoded pictures from the first encoder to implement thereon a process for a second super-resolution enlargement to create a set of super-resolution enlarged decoded pictures with a resolution higher than the standard resolution

320. The Phek-Martins-He combination would have rendered this claim element obvious for at least the reasons described above with respect to [2.3].

[3.4] a third encoder configured to have the set of super-resolution enlarged pictures from the first super-resolution enlarger as a set of encoding target pictures, employing the set of super-resolution enlarged decoded pictures from the second super-resolution enlarger as a set of reference pictures, to implement thereon a third combination of processes for a prediction and an encoding to create a third sequence of encoded bits

321. The only difference between [3.4] and [1.5] is that [3.4] only requires one set of reference pictures, namely “the set of super-resolution enlarged decoded pictures from the second super-resolution enlarger.” Thus, Phek discloses this limitation for the same reasons as described for [1.5]. *Supra* [1.5].

[3.5] wherein a spatial resolution of the set of super-resolution enlarged pictures, and that of the set of super-resolution enlarged decoded pictures are made equal

322. The only difference between [3.5] and [1.6] is that [3.5] only requires one set of reference pictures, namely “the set of super-resolution enlarged pictures” to be made equal to the spatial resolution of the “the set of super-resolution enlarged pictures.” Thus, the combination renders obvious this limitation for the same reasons as described for [1.6]. *Supra* [1.6].

[3.6] wherein the third encoder controls a selection of a set of reference pictures and creates a set of data on the selection of the set of reference pictures to identify a selected set of reference pictures during the process for prediction of the third combination

323. The Phek-Martins-He combination renders obvious this claim element for at least the reasons described above with respect to [2.5.ii].

IX. Ground 2 – Claims 1-3 are Obvious under 35 U.S.C. § 103 based on Sakazume in view of Martins and He

A. Sakazume

324. Sakazume¹⁶ discloses a video coding system where “input video signal is encoded at a plurality of coding layers exhibiting different spatial resolutions.” Ex.1046 [Sakazume], Abstract.

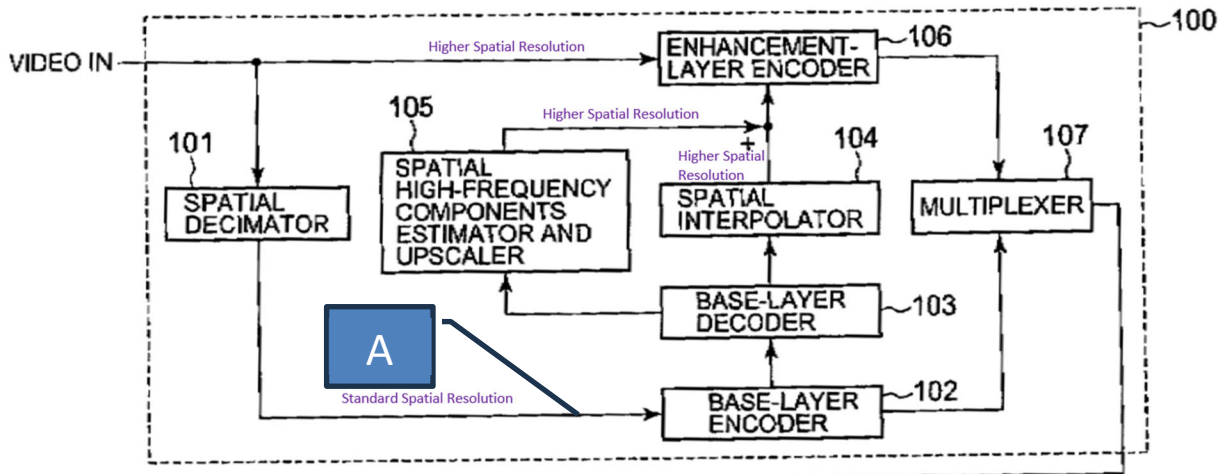
325. In the system of Sakazume, considering the base layer spatial resolution signal entering the base-layer encoder 102 – “A” in the annotated figure below –

¹⁶ The first-named inventor of Sakazume is the inventor of the challenged ’448 Patent.

“[s]patial interpolation is applied to the decoded signal of the lower coding layer to generate an upscaled decoded video signal.” *Id.*, Abstract, [0039]-[0041], Fig. 1 (reproduced below).

326. Specifically, “[t]he spatial interpolator 104 applies a spatial interpolation (scale up) procedure to the base-layer decoded video frame (locally decoded signal), thus generating an upscaled video frame (step S104).” *Id.*, [0039]. “The upscaled video frame exhibits a spatial resolution corresponds to that before the scale-down procedure at the spatial decimator 101.” *Id.*. In addition, “[a] spatial high-frequency components estimation and scale up procedure is applied to the decoded signal of the lower coding layer to generate a high-frequency components signal.” *Id.*, Abstract.¹⁷

¹⁷ The “high-frequency components signal” generated using the “high-frequency components estimation and scale up procedure” corresponds to a “super-resolution process” as detailed throughout the ’448 Patent specification. *See, e.g.*, Ex.1001 [’448 Patent], 22:51-54.



Ex. 1046 [Sakazume], FIG. 1.

327. Sakazume's Figure 1 shows that the output of spatial high-frequency components estimator and upscaler 105 and spatial interpolator 104 are added before being provided to encoder 106, but also contemplates that "the frames generated by the estimator and upscaler 105 and the interpolator 104 may be separately supplied to the encoder 106." Ex.1046 [Sakazume], [0041] (illustrated below).

328. Sakazume discloses many different approaches to providing superresolution processing. *See, e.g.*, Ex.1046 [Sakazume], 22:7-50. These include well-established signal processing techniques, such as the Maximum Likelihood (ML) method and Projection onto Convex Sets (POCS) method. These techniques have been used to process single images as well as video sequences to produce superresolution enhancement of special frequency ranges.

B. Combination of Sakazume, Martins, and He

329. Like the '448 Patent, Sakazume discloses a layered video coding system that enhances reference picture quality in an enhancement layer by use of a decoded spatially interpolated base-layer frame as well as high frequency components generated from the encoded and decoded base layer as the reference. *Supra*, §V.

330. The following subsections address straightforward modifications to Sakazume's system that would have been obvious to implement in view of Martins and He.¹⁸

1. Obvious Based on Martins to Use Super-resolution for Upscaling Base-layer Resolution Video

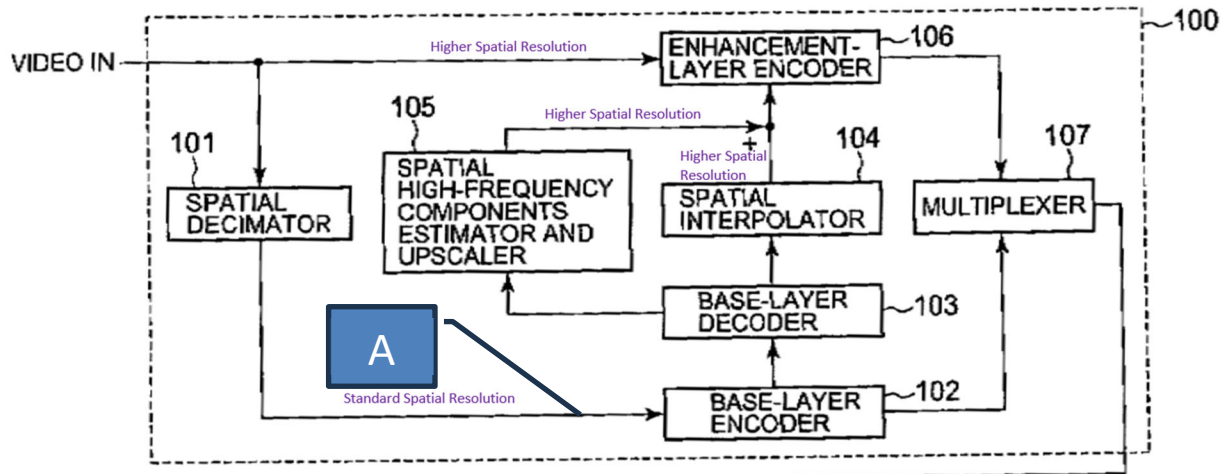
331. Sakazume describes a layered coding system with enhancement layers at higher spatial resolution than the base layer. Ex.1046 [Sakazume], [0002]-[0004]. In the system of Sakazume, with reference to Figure 1 in particular, Sakazume's apparatus "receives enhancement-layer picture element signals that carry enhancement-layer video frames" (i.e., shown in Figure 1 as Video In) which "exhibit a spatial resolution for enhancement-layer coding." Ex.1046 [Sakazume], [0036]. Sakazume then teaches that this input is decimated "producing a ... base-

¹⁸ This overview of the Sakazume-Martins-He combination is referenced and incorporated in the analysis of each claim element in Section IX.C below.

layer video frame.” Ex.1046 [Sakazume], [0036]. Although Sakazume itself describes receiving enhancement-layer video frames and decimating those frames to base-layer video frames, a POSITA would have recognized that Sakazume’s teachings would apply to situations in which base-layer video frames are received and processed (instead of processing decimated enhancement-layer video frames). And, indeed, pre-processing of input video pictures, including raw input video pictures, to enhance those input video pictures before encoding, as taught by Phek, for example. *See, e.g.*, Ex.1005 [Phek], Fig. 1a, 1b, 2.

332. Indeed, given the prior art within the knowledge of a POSITA, a POSITA would have recognized that Sakazume’s teachings are clearly applicable to not just the situation in which enhancement layer video is input, but also to situations in which the enhancement layer is generated by processing from the base layer.

333. For example, as shown in the annotated portion of Figure 1 of Sakazume included below again (and above in the summary of Sakazume), the base-layer encoder 102 receives standard spatial resolution video after higher spatial resolution video is input and decimated by spatial decimator 101, and the enhancement-layer encoder 106 receives higher spatial resolution as the input video.



Ex. 1046 [Sakazume], FIG. 1

334. A POSITA would have recognized at the time of the '448 Patent, given at least the prior art cited in my declaration, that there were other known configurations consistent with Sakazume's Figure 1 where the base-layer encoder 102 would receive standard spatial resolution video and the enhancement-layer encoder 106 would receive higher spatial resolution video.

335. One such example is in the proposed Sakazume-Martins-He combination, where base layer (i.e., lower resolution or standard spatial resolution) video is received as video input as shown in the figure below, and processed through the teachings of the Sakazume-Martins-He combination, to output higher resolution target video to the enhancement layer encoder 106. Base-layer encoder 102 receives standard spatial resolution video, consistent with Sakazume's Figure 1.

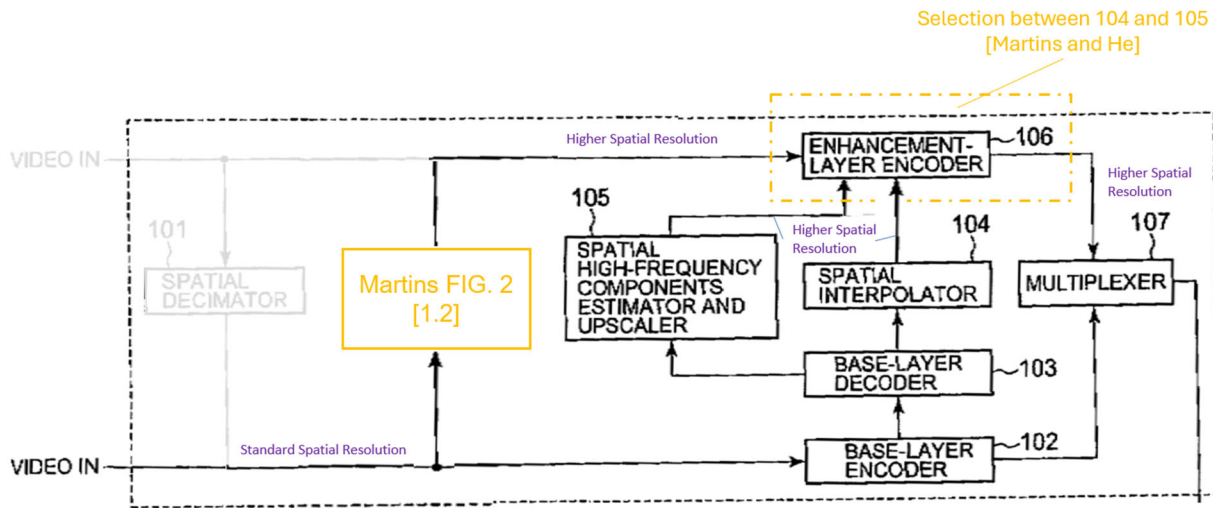


Illustration of the Sakazume-Martins-He combination.

336. A POSITA would have understood that it was obvious to generate the higher spatial resolution video as input to the enhancement-layer encoder 106, using an upsampling technique, such the super-resolution upsampling disclosed in Martins, which would be applied to a base layer sequence of video, to create a higher resolution enhancement layer sequence, as a POSITA would have known that upscaling or upsampling and decimation were known options depending on the resolution of the input stream and the desired output resolutions for the base and enhancement layers. Thus, the proposed combination of teachings above, where standard spatial resolution video is received as input and upsampled according to Martins, would have been nothing more than an obvious variation of Sakazume's teachings that was readily apparent to a POSITA considering the reference. The output of the Martins' Figure 2 super-resolution processing would provide to the

enhancement layer encoder 106 of Sakazume video frames at the same, higher resolution as in Sakazume without modifications, i.e., enhancement-layer video frames.

337. Further, this is consistent with Martins' explicitly stated goal of "improv[ing] MPEG-2 decoding, or rather to postprocess the decoded sequence re-using information in the MPEG-2 bitstream to obtain a sequence of **higher fidelity**, especially with regard to the artifacts." Ex.1007 [Martins], 1. Further, "[t]he super-resolution pictures obtained by the [Martins] algorithm are of much higher visual quality and have lower [mean-squared error] than super-resolution pictures obtained by simple spatial interpolation." Ex.1007 [Martins], 1. Thus, the Martins algorithm provides higher visual quality super-resolution pictures than the pictures output by the spatial high-frequency components estimator and upscaler 105 and spatial interpolator 104 of Sakazume, which rely on interpolation; accordingly, the higher-resolution video output by the combination would have desirable "higher visual quality".

338. Using Martin's super-resolution upsampling in the combination to enhance base layer video would have been obvious as a predictable application of known techniques to yield predictable results for which a POSITA would have had a reasonable expectation of success, especially since super-resolution upsampling was commonly performed before the '448 Patent.

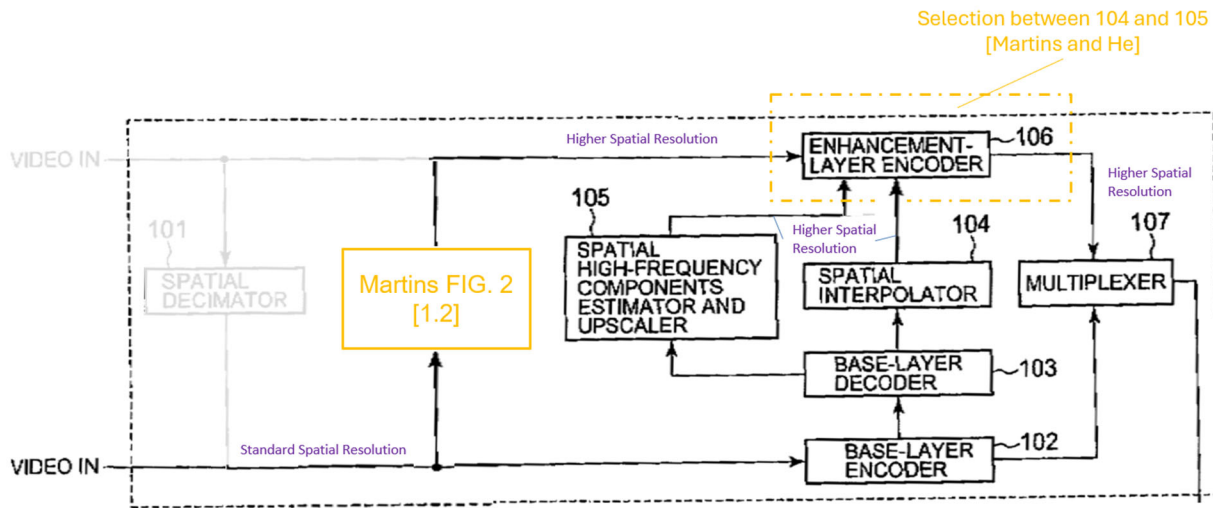


Illustration of the Sakazume-Martins-He combination.

2. Obvious Based on Martins and He to Selectively Use Super-Resolution or Interpolated Images

339. As discussed, the teachings of Sakazume-Martins render obvious techniques for performing super-resolution processing and upsampling of reference images used for inter-layer prediction of images in the enhancement layer. Ex.1046 [Sakazume], [0039], [0055], FIGS. 1-2; Ex.1007 [Martins], 1-2, 6; *supra*, §IX.B.1.

340. Sakazume also teaches applying spatial interpolation (i.e., spatial interpolator 104) and spatial high-frequency components estimating and upscaling (i.e., in unit 105) and separately providing the frames output by those units to the enhancement layer encoder as reference images. *See, e.g.*, Ex.1046 [Sakazume], ¶¶41 (“frames ... may be separately supplied”), 42-43, 57-61.

341. For example, Sakazume teaches, with respect to Figure 1, that the enhancement-layer encoder 106 performs a prediction procedure between an

enhancement-layer video frame and that “the frames generated by the estimator and upscaler 105 and the interpolator 104 may be separately supplied to the encoder 106.” Ex.1046 [Sakazume], [0041], *see also* [0043] (“and the composite base-layer high-frequency-components-estimated video frame (**or, the upscaled video frame from the spatial interpolator 104 and the high-frequency components frame from the spatial high-frequency components estimator and upscaler 105**)).”

342. In this scenario, i.e., when the frame from 104 and the frame from 105 are received separately, a POSITA would have recognized that the enhancement layer encoder 106 would select between the two frames to perform prediction according to the frame that would result in the lowest amount of residual information to reduce the amount of information needed for coding (see Technical Background). Thus, Sakazume teaches performing a selection between input frames used for prediction.

343. A POSITA would have understood that Martins and He explicitly teach such selection techniques. For example, as described with reference to Ground 1, a POSITA would have understood that Martins teaches selection on a pixel basis between super-resolution pixels obtained by motion compensation and super-

resolution pixels obtained by interpolation¹⁹ (and encoding data on the selection, *see, e.g.,* Martins’ “attribute” described at Ex.1007 [Martins], 807), which is a selection operation analogous to the selection between the outputs of Sakazume’s units 105 and 104, respectively, and He similarly teaches selective in-loop filtering using deblocking control flags. Ex.1008 [He], [0028], [0031], [0037], [0076], [0069]. Thus, combining the teachings of Martins and/or He of encoding information reflecting whether a pixel or a picture was obtained by one of two techniques would have been obvious when considering Sakazume, as doing so would have been nothing more than the combination of prior art elements according to known techniques (e.g., encoding data reflecting a selection).

344. In the combination, and based on Martins’ and He’s teaching, control flags in the frame can predictably indicate whether super-resolution processing was applied when encoding a given frame in the enhancement layer. Ex.1008 [He], [0028], [0031], [0037], [0076], [0069]. When the super-resolution processing is not applied, the reference frames are generated by resolution conversion of the base layer pictures, without super-resolution processing.

¹⁹ The ’448 Patent specification discusses performing the selection on a picture or “region” basis, where a region is defined as “... pictures each respectively divided with no spaces left into regions of a prescribed area and first and second reference pictures likewise divided with no spaces left into regions of a prescribed area”. Ex.1001 [’448 Patent], 15:41-45. Pixels meet this definition.

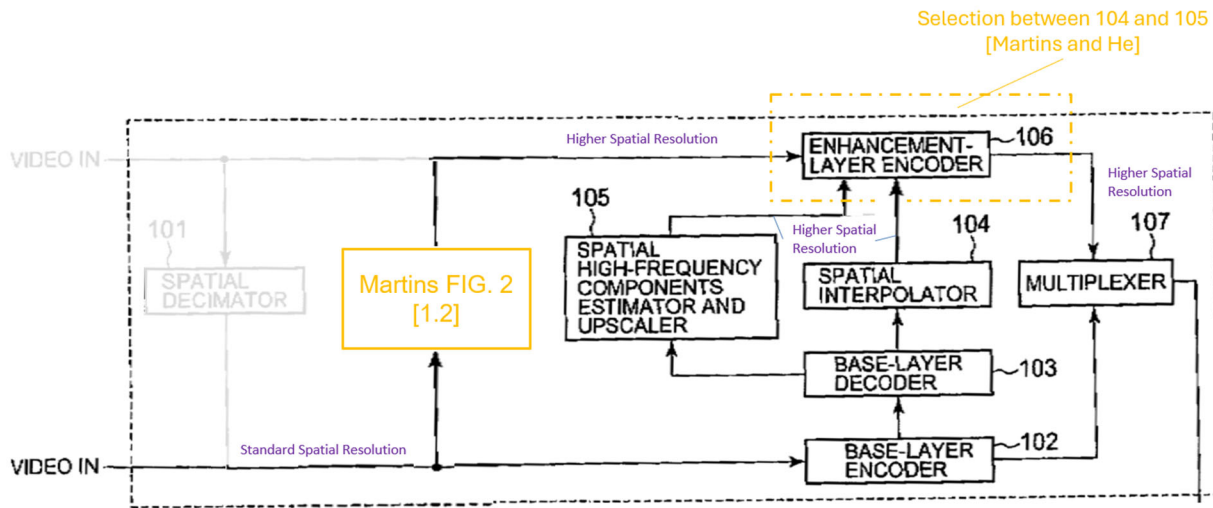


Illustration of the Sakazume-Martins-He combination.

345. The fact that selective loop filtering was standard in video coders before the '448 Patent, and the presence of such selection in Sakazume itself, would have led to a POSITA to have a reasonable expectation of success implementing data reflecting selective super-resolution processing in the combination.

C. Relevance of the Sakazume-Martins-He Combination to the Claims

346. Based on my review, the combined teachings of Sakazume in view of Martins and He as described in the preceding section would have rendered each of claims 1-3 of the '448 Patent obvious before the Critical Date.

1. Claim 1

[1.P] A moving picture encoding system that makes an encoding of a sequence of moving pictures with a resolution higher than a standard resolution using moving pictures contents which include a sequence of moving pictures with the standard resolution and do not include a sequence of moving pictures with a

resolution higher than the standard resolution, the moving picture encoding system comprising:

347. Sakazume-Martins-He renders obvious [1.P].

348. For example, Sakazume describes that its invention “relates to layered coding and decoding methods, apparatuses, and programs for moving pictures based on spatial- and time-domain and also interlayer correlation.” Ex.1046 [Sakazume], [0002]; *see also id.*, [0010], Fig. 1.

349. Sakazume discloses that “[a]n input video signal is encoded at a plurality of coding layers exhibiting different spatial resolutions.” Ex.1046 [Sakazume], [Abstract]. Specifically, “[a] spatial-domain resolution scalable layered video coding scheme produces interlayer predictive signals for use in coding enhancement layers that exhibit higher spatial resolutions than base layers also to be coded” (“*an encoding of a sequence of moving pictures with a resolution higher than a standard resolution*”). *Id.*, [0004].

350. As discussed below, the enhancement layer is encoded “*using moving pictures contents which include a sequence of moving pictures with the standard resolution and do not include a sequence of moving pictures with a resolution higher than the standard resolution.*” *Infra*, §IX.C([1.1] and [1.2]). Specifically, in the proposed combination, a POSITA would have recognized that Sakazume’s teachings were applicable to base layer video, i.e., “*moving pictures contents which*

include a sequence of moving pictures with the standard resolution and do not include a sequence of moving pictures with a resolution higher than the standard resolution” and thus the proposed combination renders obvious the preamble of claim 1.

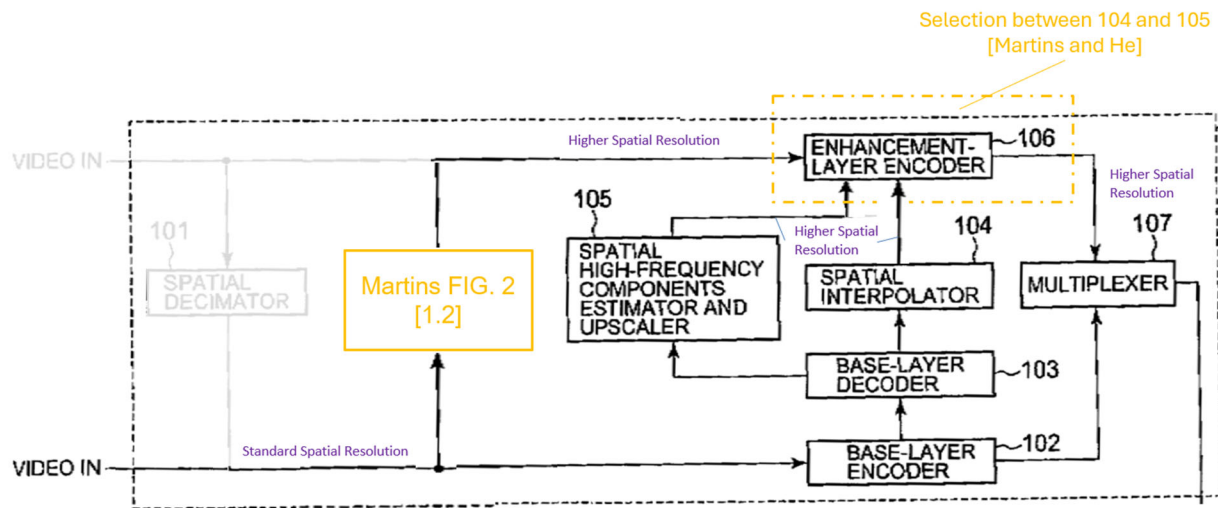


Illustration of the Sakazume-Martins-He combination.

[1.1] a first encoder configured to work on a sequence of moving pictures with a standard resolution to implement a first combination of processes for an encoding and a decoding to create a first sequence of encoded bits and a set of decoded pictures with the standard resolution

351. Sakazume-Martins-He renders obvious [1.1].

352. Sakazume describes a “base-layer encoder 102” and a “base-layer decoder 103” that together are the claimed “*first encoder*.” Ex.1046 [Sakazume], [0036]-[0038]; Fig. 1.

353. In the proposed Sakazume-Martins-He combination, the “base-layer encoder 102” receives and processes “base-layer video frame[s]” (i.e., standard

spatial resolution video) to generate a “base-layer bitstream” (“*a first sequence of encoded bits*”) and sends this to a “base-layer decoder 103.” *Id.*, [0036]-[0038]; Fig.

1. In the Sakazume-Martins-He combination, the resolution of the base-layer is the claimed “standard resolution.” *Id.*, [0004].

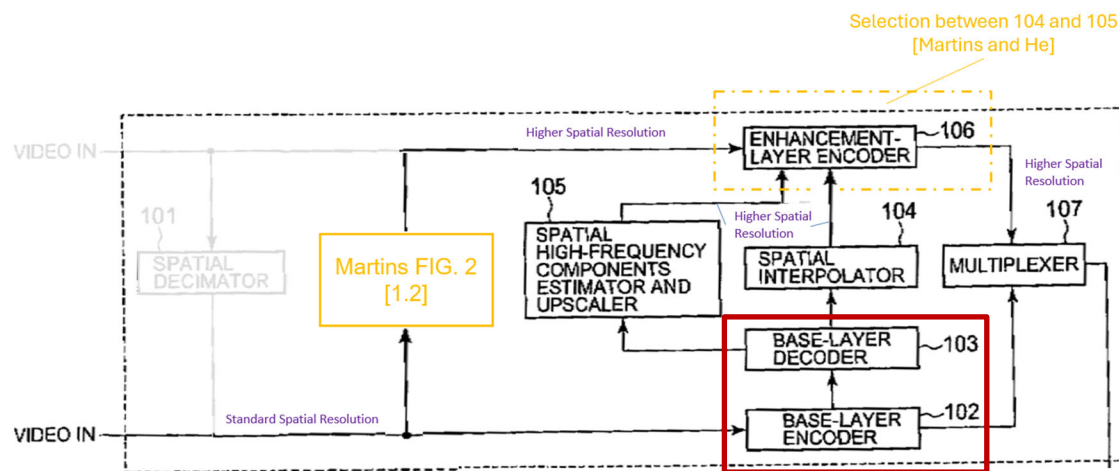


Illustration of the Sakazume-Martins-He combination.

354. Sakazume further teaches that “base-layer decoder 103 decodes the base-layer bitstream generated by the base-layer encoder 102 (step 5103), thus generating a base-layer decoded video frame” (“*set of decoded pictures with the standard resolution*”). *Id.*, [0038].

[1.2] a first super-resolution enlarger configured to work on the sequence of moving pictures with the standard resolution to implement a process for a first super-resolution enlargement to create a set of super-resolution enlarged pictures with a resolution higher than the standard resolution

355. Sakazume-Martins-He renders obvious [1.2].

356. As detailed above, it would have been obvious based on the teachings of Martins, that when the input moving picture video signal is base layer video frames, instead of inputting enhancement layer video frames, the base layer video frames could be upsampled by super-resolution processing to create enhancement layer video frames. *Supra*, §V.A.2.a.

357. Thus, in the combination, as illustrated below, in addition to the encoding/decoding by Sakazume's base layer encoder 102 and base layer decoder 103, Martins' super-resolution upsampling technique ("***a first super-resolution enlarger***") is applied to the input moving picture video signal at base layer spatial resolution ("***subsequence of the sequence of moving pictures with the standard resolution***") to create enhancement layer video frames at a higher spatial resolution ("***set of super-resolution enlarged pictures with a first resolution higher than the standard resolution***"). *Id.*

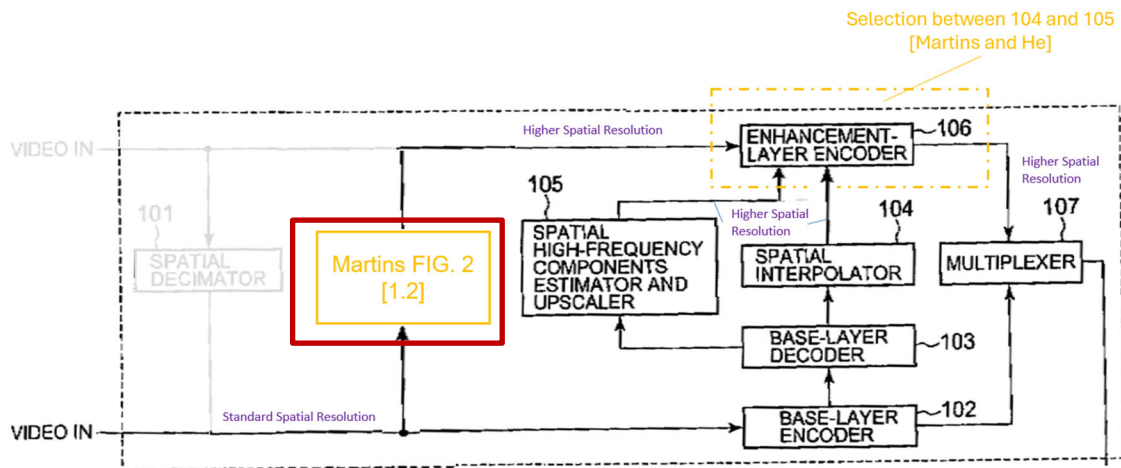


Illustration of the Sakazume-Martins-He combination.

[1.3] a second super-resolution enlarger configured to acquire the set of decoded pictures from the first encoder to implement thereon a process for a second super-resolution enlargement to create a set of super-resolution enlarged decoded pictures with a resolution higher than the standard resolution

358. Sakazume-Martins-He renders obvious [1.3].

359. Sakazume discloses that “spatial high-frequency components estimator and upscaler 105” (“**a second super-resolution enlarger**”) “applies a spatial high-frequency components estimation and scale-up procedure, in a specific direction in the spatial domain, to the base-layer decoded frame supplied by the base-layer decoder 103” i.e., the estimator and upscaler 105 “**acquire[s] the set of decoded pictures from the first encoder to implement thereon a process for a second super-resolution enlargement**” as recited. Ex.1046 [Sakazume], [0040]. Sakazume further discloses that “spatial high-frequency components estimator and upscaler 105 “generates a **high frequency components frame** that carries estimated high

frequency components. The estimated high-frequency components correspond to a video frame that exhibits a spatial resolution before subjected to the scale-down procedure” (Ex.1046 [Sakazume], [0040]), i.e., the high frequency components frame is a picture “*with a resolution higher than the standard resolution*” as recited. *Supra*, §IX.A.

360. Sakazume’s “estimator and upscaler 105” corresponds to a “*second super-resolution enlarger*” given the ’448 Patent’s broad description that “the process for the prescribed super-resolution enlargement may for instance be an application of **any such super-resolution process as described above**” (Ex. 1001 [’448 Patent], 22:51-54, *see also* 22:4-50) which description would encompass techniques such as those described in Sakazume. For example, Sakazume’s “spatial high-frequency components estimator and upscaler 105 applies a spatial high-frequency components estimation and scale-up procedure, in a specific direction in the spatial domain.” Ex.1046 [Sakazume], [0040]. This is similar to, for example, the technique described in Park (Ex.1012), which is described in the ’448 Patent at Ex.1001 [’448 Patent], 22:7-16. Thus, Sakazume’s high-frequency components estimation and scale-up procedure corresponds to “*super-resolution enlargement*” as recited in this limitation.

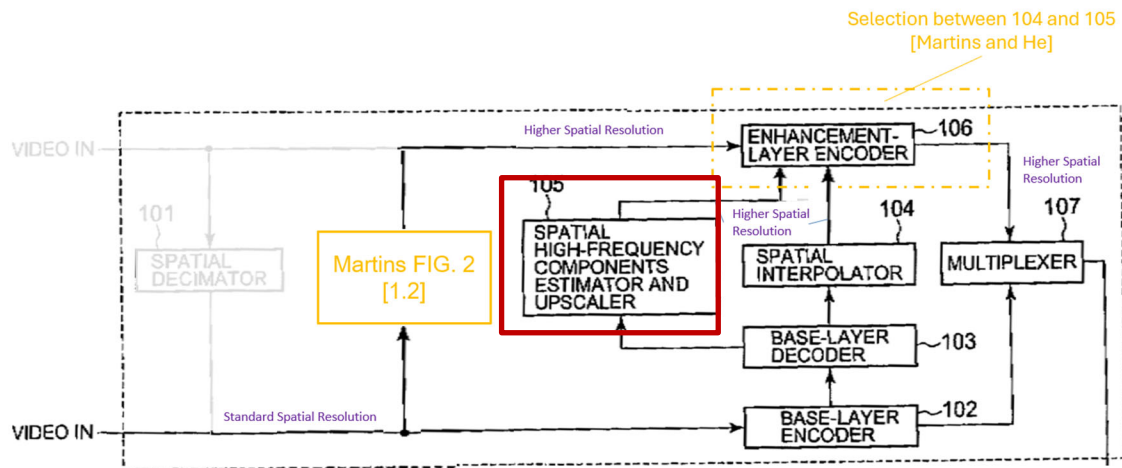


Illustration of the Sakazume-Martins-He combination.

[1.4] a third resolution converter configured to acquire the set of decoded pictures from the first encoder to implement thereon a process for a third resolution conversion to create a set of resolution converted enlarged decoded pictures with a resolution higher than the standard resolution

361. Sakazume-Martins-He renders obvious [1.4].

362. Sakazume discloses that “[t]he spatial interpolator 104 applies a spatial interpolation (scale up) procedure to the base-layer decoded video frame (locally decoded signal), thus generating an upscaled video frame (step S104)” and that “[t]he upscaled video frame exhibits a spatial resolution correspond[ing] to ... the **spatial resolution of the enhancement layer** which is one layer higher than the base layer.” Ex.1046 [Sakazume], [0039].

363. Thus, Sakazume’s “spatial interpolator 104” corresponds to the recited **“third resolution converter”** and, as it receives the “base-layer decoded video” frame and converts it to a spatial resolution of the enhancement layer, the

interpolator 104 “*implement[s] thereon a process for a third resolution conversion to create a set of resolution converted enlarged decoded pictures with a resolution higher than the standard resolution*” as recited.

364. As detailed below, the teachings of Martins and He describe selecting between two pictures, one being a super-resolution enlarged picture (e.g., the output of Sakazume’s estimator and upscaler 105) and the other being an interpolated picture (e.g., the output of Sakazume’s interpolator 104). *Supra*, §IX.B.2.

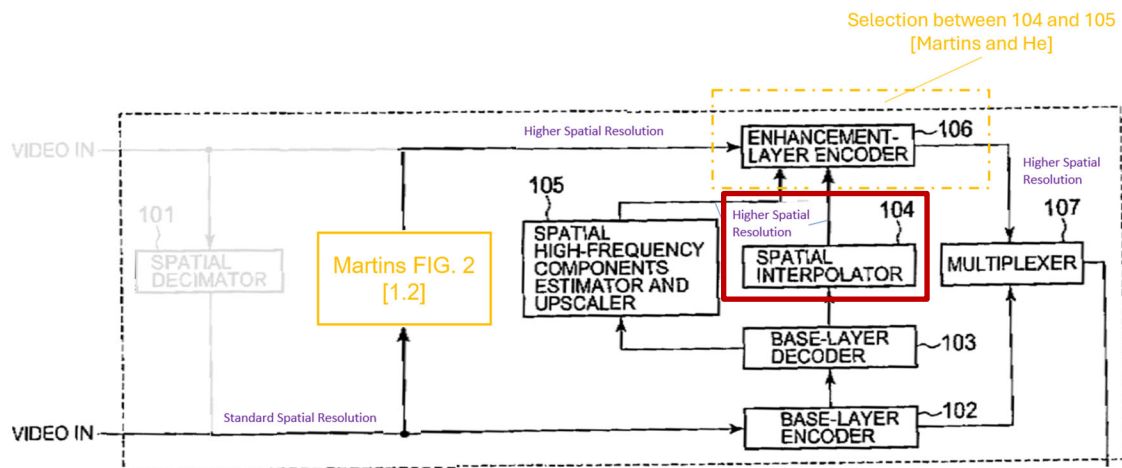


Illustration of the Sakazume-Martins-He combination.

[1.5] a third encoder configured to have the set of super-resolution enlarged pictures from the first super-resolution enlarger as a set of encoding target pictures, employing the set of super-resolution enlarged decoded pictures from the second super-resolution enlarger and the set of resolution converted enlarged decoded pictures from the third resolution converter as sets of reference pictures, to implement thereon a third combination of processes for a prediction and an encoding to create a third sequence of encoded bits

365. Sakazume-Martins-He renders obvious [1.5].

366. Sakazume discloses that the enhancement-layer encoder 106 (“***third encoder***”) receives enhancement-layer video frames from the input moving-picture video signal, processed by Martins Fig. 2 resulting in higher spatial resolution video, as encoding targets and uses upscaled video frames from the spatial interpolator 104 as reference frames in order to perform an enhancement-layer coding procedure and create an enhancement layer bitstream (“***implement thereon a third combination of processes for a prediction and an encoding to create a third sequence of encoded bits***”). Ex.1046 [Sakazume], [0042]-[0043], [0045].

367. As discussed above, in the Sakazume-Martins-He combination, the enhancement-layer encoder 106 (“***third encoder***”) receives enhancement-layer video frames created by estimation and upscaling of the standard resolution input moving-picture video signal from the first super-resolution enlarger (the Martins’ Figure 2 super-resolution processing) (“the ***set of super-resolution enlarged pictures from the first super-resolution enlarger***”) as encoding targets, super-resolution enlarged decoded pictures from the second super-resolution enlarger (the Sakazume spatial high-frequency components estimator and upscaler 105) (a second “***set of super-resolution enlarged decoded pictures***”) as a first set of reference pictures, and resolution converted decoded pictures from the resolution converter (the Sakazume spatial interpolator 104) (“***set of resolution converted enlarged decoded pictures***”) as a second set of reference pictures. *Supra*, §IX.C([1.2] to

[1.4]); Ex.1046 [Sakazume], ¶¶41 (“Or, the frames generated by the estimator and upscaler 105 and the interpolator 104 may be separately supplied to the encoder 106.”), 42-43, 57-61 (describing providing output of units 105 and 104 separately).

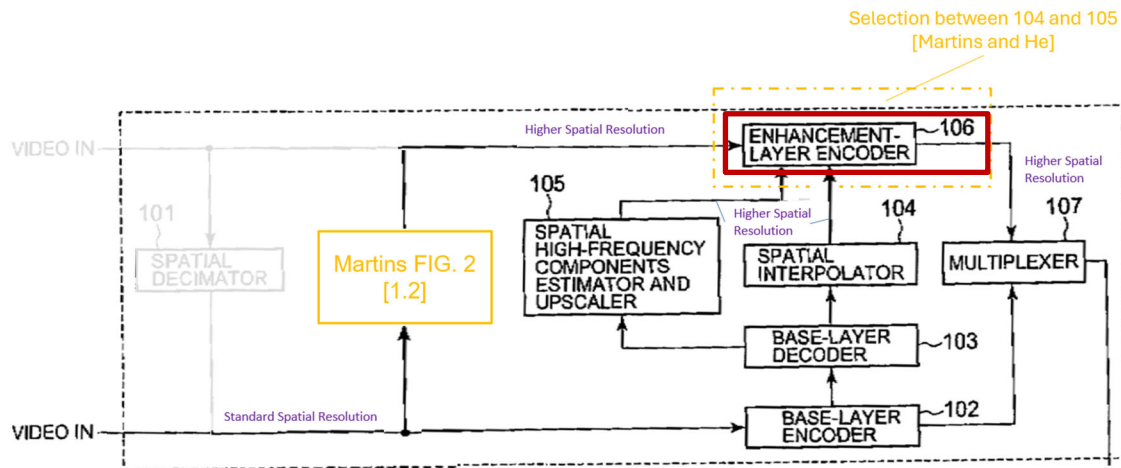


Illustration of the Sakazume-Martins-He combination.

[1.6] wherein a spatial resolution of the set of super-resolution enlarged pictures, that of the set of super-resolution enlarged decoded pictures, and that of the set of resolution converted enlarged decoded pictures are made equal

368. Sakazume-Martins-He renders obvious [1.6].

369. As discussed, Sakazume discloses further that “[t]he upscaled video frame exhibits a spatial resolution correspond[ing] to ... the spatial resolution of the enhancement layer which is one layer higher than the base layer.” Ex.1046 [Sakazume], [0039], [0042]-[0043], [0045], FIGs. 1-2; *supra*, §V.B.4.

370. In the Sakazume-Martins-He combination, as described above, the first set of reference pictures is provided by the second super-resolution enlarger (i.e.,

Sakazume's spatial high-frequency components estimator and upscaler 105) and the second set of reference pictures is provided by the third resolution converter (i.e., Sakazume's spatial interpolator 104). Based on Sakazume's teachings, it would have been obvious to convert not only the first set of reference pictures, but also the second set of reference pictures, to the same spatial resolution as the enhancement layer. And a POSITA would have been motivated to do so with a reasonable expectation of success for the same reasons as described for Ground 1. *Supra*, §§II.B.G.5-6; VIII.A.2. Sakazume explicitly teaches that the "upscaled video frame" from "spatial interpolator 104" "exhibits a spatial resolution [that] corresponds to ...the spatial resolution of the enhancement layer" and likewise, the "estimated high-frequency components" from "estimator and upscaler 105" "correspond to a video frame that exhibits a spatial resolution before subjected to the scale-down procedure at the spatial decimator 101" (Ex.1046 [Sakazume], [0039-0040]); that is, the resolution of the output of 104 and 105 in Sakazume are "*made equal*" as recited.

[1.7] wherein the third encoder controls a selection of a set of reference pictures and creates a set of data on the selection of the set of reference pictures to identify a selected set of reference pictures during the process for prediction of the third combination

371. Sakazume-Martins-He renders obvious [1.7].

372. The Sakazume-Martins-He combination's enhancement layer encoder 106 ("***third encoder***") encodes pictures in the enhancement layer bitstream by controlling selection of either decoded base layer pictures enhanced by super-resolution enlargement (e.g., pictures output by Sakazume's spatial high-frequency components estimator and upscaler 105) or decoded base layer pictures not enhanced by SR-U filtering (e.g., pictures output by Sakazume's spatial interpolator 104) as reference pictures. It would have been obvious based on Martins, as well as He, to use He's teachings of a "deblocking control flag" as "***a set of data on the selection of the set of reference pictures to identify a selected set of reference pictures***" for the same reasons as described for Ground 1. *Supra*, § II.G.F; IV.B.8.

2. Claim 2

[2.P] A moving picture encoding system that makes an encoding of a sequence of moving pictures with a resolution higher than a standard resolution using moving pictures contents which include a sequence of moving pictures with the standard resolution and do not include a sequence of moving pictures with a resolution higher than the standard resolution, the moving picture encoding system comprising:

373. The Sakazume-Martins-He combination would have rendered this claim element obvious for at least the reasons described above with respect to [1.P].

[2.1] a first encoder configured to work on a sequence of moving pictures with a standard resolution to implement a first combination of processes for an encoding and a decoding to create a first sequence of encoded bits and a set of decoded pictures with the standard resolution

374. The Sakazume-Martins-He combination would have rendered this claim element obvious for at least the reasons described above with respect to [1.1].

[2.2] a first super-resolution enlarger configured to work on the sequence of moving pictures with the standard resolution to implement a process for a first super-resolution enlargement to create a set of super-resolution enlarged pictures with a resolution higher than the standard resolution

375. The Sakazume-Martins-He combination would have rendered this claim element obvious for at least the reasons described above with respect to [1.2].

[2.3] a third resolution converter configured to acquire the set of decoded pictures from the first encoder to implement thereon a process for a third resolution conversion to create a set of resolution converted enlarged decoded pictures with a resolution higher than the standard resolution

376. The Sakazume-Martins-He combination would have rendered this claim element obvious for at least the reasons described above with respect to [1.3].

[2.4] a third encoder configured to have the set of super-resolution enlarged pictures from the first super-resolution enlarger as a set of encoding target pictures, employing the set of resolution converted enlarged decoded pictures from the third resolution converter as a set of reference pictures, to implement thereon a third combination of processes for a prediction and an encoding to create a third sequence of encoded bits

377. The only difference between [2.4] and [1.5] is that [2.4] only requires one set of reference pictures, namely “the set of resolution converted enlarged decoded pictures from the third resolution converter.” Thus, the combination renders obvious this limitation for the same reasons as described for [1.5]. *Supra* [1.5].

[2.5] wherein a spatial resolution of the set of super-resolution enlarged pictures, and that of the set of resolution converted enlarged decoded pictures are made equal

378. The only difference between [2.5] and [1.6] is that [2.5] only requires one set of reference pictures, namely “the set of resolution converted enlarged decoded pictures” to be made equal to the spatial resolution of the “the set of super-resolution enlarged pictures.” Thus, the combination renders obvious this limitation for the same reasons as described for [1.6]. *Supra* [1.6].

[2.6] wherein the third encoder controls a selection of a set of reference pictures and creates a set of data on the selection of the set of reference pictures to identify a selected set of reference pictures during the process for prediction of the third combination

379. The Sakazume-Martins-He combination would have rendered this claim element obvious for at least the reasons described above with respect to [1.7].

3. Claim 3

[3.P] A moving picture encoding system that makes an encoding of a sequence of moving pictures with a resolution higher than a standard resolution using moving pictures contents which include a sequence of moving pictures with the standard resolution and do not include a sequence of moving pictures with a resolution higher than the standard resolution, the moving picture encoding system comprising

380. The Sakazume-Martins-He combination would have rendered this claim element obvious for at least the reasons described above with respect to [1.P].

[3.1] a first encoder configured to work on a sequence of moving pictures with a standard resolution to implement a first combination of processes for an encoding and a decoding to create a first sequence of encoded bits and a set of decoded pictures with the standard resolution

381. The Sakazume-Martins-He combination would have rendered this claim element obvious for at least the reasons described above with respect to [1.1].

[3.2] a first super-resolution enlarger configured to work on the sequence of moving pictures with the standard resolution to implement a process for a first super-resolution enlargement to create a set of super-resolution enlarged pictures with a resolution higher than the standard resolution

382. The Sakazume-Martins-He combination would have rendered this claim element obvious for at least the reasons described above with respect to [1.2].

[3.3] a second super-resolution enlarger configured to acquire the set of decoded pictures from the first encoder to implement thereon a process for a second super-resolution enlargement to create a set of super-resolution enlarged decoded pictures with a resolution higher than the standard resolution

383. The Sakazume-Martins-He combination would have rendered this claim element obvious for at least the reasons described above with respect to [1.3].

[3.4] a third encoder configured to have the set of super-resolution enlarged pictures from the first super-resolution enlarger as a set of encoding target pictures, employing the set of super-resolution enlarged decoded pictures from the second super-resolution enlarger as a set of reference pictures, to implement thereon a third combination of processes for a prediction and an encoding to create a third sequence of encoded bits

384. The only difference between [3.4] and [1.5] is that [3.4] only requires one set of reference pictures, namely “the set of super-resolution enlarged decoded pictures from the second super-resolution enlarger.” Thus, the combination renders obvious this limitation for the same reasons as described for [1.5]. *Supra* [1.5].

[3.5] wherein a spatial resolution of the set of super-resolution enlarged pictures, and that of the set of super-resolution enlarged decoded pictures are made equal

385. The only difference between [3.5] and [1.6] is that [3.5] only requires one set of reference pictures, namely “the set of super-resolution enlarged pictures” to be made equal to the spatial resolution of the “the set of super-resolution enlarged pictures.” Thus, the combination renders obvious this limitation for the same reasons as described for [1.6]. *Supra* [1.6].

[3.6] wherein the third encoder controls a selection of a set of reference pictures and creates a set of data on the selection of the set of reference pictures to identify a selected set of reference pictures during the process for prediction of the third combination

386. The Sakazume-Martins-He combination would have rendered this claim element obvious for at least the reasons described above with respect to [1.7].

X. Additional Remarks

I currently hold the opinions set expressed in this declaration. But my analysis may continue, and I may acquire additional information and/or attain supplemental insights that may result in added observations.

EXHIBIT A

Cliff Reader, Ph.D.

Expertise

- Real-time Processing and Display
 - Video coding (MPEG-HEVC)
 - Audio coding (MPEG-AAC)
 - Imaging/Video Systems Architecture
 - Imaging/Video Chip Architecture
 - Videoconferencing
 - Speech Processing
 - Consumer audio/video
 - Reconnaissance Imaging
 - Medical Imaging
-

Professional Summary

Dr. Reader has over forty five years of work experience in digital imaging, digital video & audio and digital speech processing. In the 1970's and early 1980's, he was one of the leaders in the digital imaging field, performing extensive work in areas of compression, real-time processing, real-time display, image enhancement, and others. Applications included reconnaissance imaging, medical imaging, earth resources management and videoconferencing. This work anticipated much of the technology being used in contemporary consumer digital video systems. He has in-depth experience of the commercial history and when key parts of technology were first developed and published.

Dr. Reader's career includes technical work in areas of algorithm design, system design, and semiconductor chip design; and business development, marketing and sales work in areas of market analysis, product roadmap development and direct sales.

Dr. Reader has held many leadership roles in standards development: US Head of Delegation to MPEG (1991-1993); chief editor of the MPEG1 standard (1991-1993); instigator of the successful effort to make MPEG2 the market standard for digital TV over proprietary competitors in the US and national competitors in Europe and Japan (1992); MPEG4 Subcommittee Chairman (1993-1996); AVS (China) Subgroup Chairman (2003-Present). He is an expert in all major audio-video standards. MPEG audio (Layer I, II, mp3, AAC), Dolby AC3, MPEG video (1, 2, 4), H.264, Windows Media 9 (VC-1), H.265 (HEVC), AVS and AV1. He is also an expert in DVD standards, including formatting, navigation and authoring.

Dr. Reader has extensive experience in patent licensing and patent pools. He was the technical expert for developing the initial list of potentially essential patents for the first MPEGLA patent pool (MPEG2 video and systems). He is the Co-Director leading the pool formation and license negotiations for the AVS patent pools. Dr. Reader has experience in the valuation of patents and the acquisition and sale of patent portfolios.

Dr. Reader has authored over twenty technical papers, two book chapters and two patents in digital video processing, display, and compression. He taught digital signal processing and image processing at Santa Clara University in the early 1980s and is an Adjunct Professor at Peking University, Beijing, China.

Employment History

From: 2001 **Independent Consultant**
To: Present Escondido, CA
Provide technical and marketing consulting services in the areas of digital imaging, digital video, digital audio and speech including:

- consumer video
- real-time processing and display
- image and video & audio compression (specializing in standards)
- imaging/video systems architecture
- imaging/video chip architecture
- audio/video transmission, including optical & wireless networks

Over 80 clients including Apple, AT&T, Adobe, Broadcom, China Central TV, Chinese Academy of Sciences, Cisco, Comcast, DirecTV, Dish Network, Disney, EchoStar, ETRI, Facebook, Fujifilm, Gateway, GE, Google, Haier, Hisense, HP, Hulu, IBM, Intel, Johnson Controls, KAIST, Lenovo, LG, Matsushita, MediaTek, Microsoft, Motorola, Netflix, Nikon, Nokia, ON2, OPPO, Peloton, Philips, Quanta, Riverbed, Samsung, Sharp, Siemens, Sony, Starz, Sun Microsystems, Toshiba, TPV, Tyco, US Department of Justice, Verizon, Vimeo, Vizio, Wistron, Xiaomi, Yahoo, Zoran and ZTE.

Personal database of patent analysis:

- Evaluation of all the US patents in the MPEG2, H.264 and H.265 patent pools.
- Evaluation of the US patents and published applications underlying the H.264 and H.265 standards.
- Personal archive of over 2000 audiovisual patents.

Personal database of audio and video prior art.

- Virtually complete collection of all contributions to the H.26x and MPEGx families of standards from 1985 to today.
- Many hundreds of technical journal and conference articles documenting the first invention of key technologies.
- Work-in-progress book on the history of video compression.
- Historical collection of technical data from the earliest digital video, image processing and real-time image display systems in the 1970s and 1980s.

From: 1999 **nDSP Corporation**
To: 2001 Campbell, CA
Title: *Vice President of Marketing and Acting Vice President of Sales*

Responsible for marketing and sales in a startup company providing high-quality video signal processors to TV manufacturers. Developed product specifications and roadmap. Built and managed sales and marketing team of 6 direct and contractor staff in the US and Japan to provide technical sales, sales support, application engineering, Marcomm and PR. Established distribution in China, Japan and throughout the Far East. Secured first sales as company moved into production phase with design wins in China, Japan and Europe. Initiated and secured major design wins in emerging Progressive DVD business, with production volumes beginning in 4Q00 at an annualized rate of \$3.5M. Established strategic partnership with Pixelworks in LCD chip business. Pixelworks acquired nDSP in 2001.

From: 1997 **CagEnt Technologies, Inc. (wholly owned by Samsung)**
 To: 1998 Santa Clara, CA
 Position: *Vice President of Sales, Digital Video Products*
 Responsible for a line of PC-based real-time MPEG2 and Dolby Digital encoders. Established distribution agreements for third-party DVD authoring software. Secured OEM contracts for the products. Built the sales channel worldwide. Built sales to break-even point in 9 months leading to sale of the product line.

From: 1993 **Samsung Semiconductor, Inc.**
 To: 1997 San Jose, CA
 Position: *Director of Marketing, Digital Media Products;*
Acting Director of Software Engineering
 Developed business plans for digital audiovisual semiconductor chips. Managed team of over 50 software engineers in US and Korea for development of a media processor. Established strategic partnerships with Microsoft and IBM for media processor

From: 1990 **Cypress Semiconductor, Inc.**
 To: 1993 San Jose, CA
 Position: *Business Development Manager*
 Founding member of new business unit chartered to develop a “video DSP”. Jointly developed a complete business plan, and personally initiated and developed relationships with Philips, Sun Microsystems, Apple Computers and IBM. Chief architect for this early media DSP design, capable of complete MPEG1 decode in real time. (Fully validated by cycle-accurate simulation). Managed R&D team of 10 hardware and software engineers.

Head of US Delegation to MPEG, 1991 to 1992. Chief Editor of the MPEG1 standard. Technical Expert for establishing the

MPEGLA Patent Licensing Pool.

From: 1988 **Sun Microsystems, Inc.**
To: 1990 Mountain View, CA
Position: *Program Manager*
Developed requirements and specifications for real-time imaging system. Managed a team of 20 engineers to develop the architecture and high-level design. Interfaced with two key customers – Philips Medical Systems and Lockheed. Developed design for single-board real-time medical and reconnaissance image display.

From: 1981 **International Imaging Systems, Zoran, Virtual Imaging**
To: 1988 Design and development of real-time image processing and display systems. Details available on request.

From: 1975 **Ford Aerospace & ESL (A Division of TRW)**
To: 1981 Details available on request.

Consulting History

Date 2023 **Intel**
To: Present
Duties: Consulting Expert

Date 2023 **Disney**
To: Present
Duties: Consulting Expert

Date 2023 **Samsung**
To: Present
Duties: Consultant

Date 2023 **Peloton**
To: Present
Duties: Consulting Expert in NEC matter

Date 2023 **Alps Alpine**
To: 2023
Duties: Consulting Expert in DivX matter

Date 2023 **Fish & Richardson**
To: Present
Duties: Consultant

Date	2023	Oppo
To:	Present	
Duties:	Consulting Expert on speech processing in Nokia matter	
Date	2022	Samsung
To:	2023	
Duties:	Testifying Expert on speech processing in Jawbone matter	
Date	2022	Asus, Lenovo, MSI
To:	2023	
Duties:	Consulting Expert in VideoLabs matter	
Date	2022	Google
To:	2023	
Duties:	Consulting Expert on speech processing in Jawbone matter	
Date	2022	Starz
To:	Present	
Duties:	Consulting Expert in VideoLabs matter	
Date	2022	Meta Platforms (Facebook)
To:	Present	
Duties:	Consulting Expert in VideShare matter	
Date	2022	Oppo
To:	Present	
Duties:	Consulting Expert in InterDigital matter	
Date	2022	Apple
To:	2023	
Duties:	Consulting Expert on speech processing in Jawbone matter	
Date	2021	Verkada
To:	2022	
Duties:	Testifying Expert in Motorola Solutions matter	
Date	2021	Facebook
To:	2022	
Duties:	Consulting Expert in EyesMatch matter	
Date	2021	Videoconferencing company
To:	2023	
Duties:	Testifying Expert	
Date	2021	Canon

To:	2022	
Duties:		Consulting Expert in WSOU Investments/Brazos Licensing matter.
Date	2021	Axis
Duties:		Consulting Expert
Date	2021	nVIDIA
To:	2022	
Duties:		Consultant
Date	2020	Xiaomi
To:	2021	
Duties:		Consulting Expert in InterDigital matter
Date	2020	LG
To:	2021	
Duties:		Consulting Expert DivX matter
Date	2020	Vizio
To:	2023	
Duties:		Consulting Expert VDPP matter
Date	2020	Netflix
To:	2022	
Duties:		Consultant
Date	2020	Oppo
To:	Present	
Duties:		Consultant
Date	2020	IdeaHub
To:	Present	
Duties:		Preparation of declaration for Inter Partes Review proceedings. (Unified Patents matter)
Date	2019	Google
To:	2021	
Duties:		Preparation of declarations for Inter Partes Review proceedings (PMC matter)
Date	2019	Google
To:	2021	
Duties:		Consulting Expert (Multiple projects)
Date	2019	LG
To:	2020	

	Duties:	Preparation of declarations for Inter Partes Review proceedings. (Uniloc matter)
Date	2019	Klarquist, Sparkman
To:	2021	
	Duties:	Consulting Expert (Multiple projects)
Date	2019	Netflix, Hulu
To:	2021	
	Duties:	Preparation of declarations for Inter Partes Review proceedings. (DivX matter)
Date	2019	MediaTek
	Duties:	Consulting Expert in American Patents matter
Date	2019	AT&T
	Duties:	Consulting Expert
Date	2019	KAIST/KBS
To:	2021	
	Duties:	Consulting Expert for Inter Partes Review proceedings. (Unified Patents matter)
Date	2019	Vimeo
To:	2020	
	Duties:	Preparation of declarations for Inter Partes Review proceedings. (BT matter)
Date	2019	Apple
	Duties:	Preparation of declarations for Inter Partes Review proceedings. (Red.com matter)
Date	2019	Adobe
To:	2021	
	Duties:	Preparation of declarations for Inter Partes Review proceedings and consulting expert. (Realtime matter)
Date	2018	Toshiba
To:	2019	
	Duties:	Testifying expert concerning valuation of essential patents for ATSC standards
Date	2018	Hisense
	Duties:	Analysis of essential patents for AAC standards
Date	2018	HTC

To:	2020	
Duties:		Analysis of essential patents for the H.265 standard
Date	2018	Cisco
To:	2020	
Duties:		Consulting expert in Realtime matter
Date	2017	Dish Network
To:	2023	
Duties:		Consulting expert in ClearPlay matter
Date	2017	Apple
To:	2019	
Duties:		Preparation of declarations for Inter Partes Review proceedings. (Uniloc matter)
Date	2017	Dish Network
To:	2020	
Duties:		Testifying expert in Realtime matter
Date	2017	CustomPlay
To:	2020	
Duties:		Preparation of declaration for summary judgment in Amazon case.
Date	2017	Vizio, Funai, LG, MediaTek, MStar
Duties:		Testifying expert in Broadcom case.
Date	2017	Apple
Duties:		Preparation of declarations for Inter Partes Review proceedings. (Nokia matter)
Date	2017	Vizio
Duties:		Preparation of declarations for Inter Partes Review proceeding. (Lightside Technologies matter)
Date	2016	Dell, Riverbed, EchoStar, Dish Network
To:	2019	
Duties:		Testifying expert in Realtime matter
Date	2016	Sharp, Vizio, ON
To:	2018	
Duties:		Testifying expert in WiLAN matter
Date	2016	Avid
To:	2017	
Duties:		Testifying expert in Harmonic matter.

Date	2016	Apple
Duties:		Testifying expert in PUMA matter.
Date	2016	Vizio
Duties:		Testifying expert in PMC matter.
Date	2016	US Dept. of Justice, Dept. of Homeland Security, DoD, FBI, NSA, & Others
To:	2021	
Duties:		Consulting expert in Discovery Patents (3 rd Eye Surveillance) matter. Declarations in support of six Inter Partes Reviews.
Date	2015	ST Micro
To:	2016	
Duties:		Consulting expert in Avago matter.
Date	2015	AT&T
Duties:		Consulting expert in Raniere matter.
Date	2015	Motorola
To:	2016	
Duties:		Consulting expert in PUMA matter.
Date	2015	TPV
To:	2016	
Duties:		Consulting expert in Thomson licensing matter.
Date	2015	Tyco (Exacq, Johnson Controls)
To:	2019	
Duties:		Consulting expert in JDS Technologies matter.
Date	2015	Tyco
Duties:		Declaration in support of Inter Partes Review petition in Hawk Technology matter.
Date	2015	Sceptre
Duties:		Expert reports in licensing matter concerning MPEGLA MPEG2 pool patents.
Date	2015	Samsung
Duties:		Consultant and fact witness in case asserting Samsung patents of which I am a named inventor, against NVidia.
Date	2015	Nissim
To:	2016	

	Duties:	Consultant expert for patent holder in 20 th Century Fox, MGM, and Paramount cases concerning patents related to DVD navigation. Testified at Markman hearings.
Date	2015	Nikon
To:	Present	
	Duties:	Consultant.
Date	2015	Humax
	Duties:	Search and analysis of prior art.
Date	2015	Apple
	Duties:	Analysis of patent portfolio potentially essential for the H.264 standard.
Date	2014	Haier
To:	2018	
	Duties:	Analysis of patent portfolios for smart consumer devices.
Date	2014	Cisco
To:	2021	
	Duties:	Consultant.
Date	2014	ZTE, Sony, LG, Samsung, Nokia, GoPro
To:	2015	
	Duties:	Consulting expert.
Date	2014	Tyco
To:	2015	
	Duties:	Expert reports; declaration in support of Inter Partes Review in Trover matter.
Date	2013	Humax
	Duties:	Consultant.
Date	2014	GE Licensing
To:	2015	
	Duties:	Consultant.
Date	2014	HikVision & Axis
To:	2015	
	Duties:	Declaration in support of Inter Partes Review. Expert reports in Trover matter.
Date	2014	Sharp
To:	2016	
	Duties:	Declaration in support of Inter Partes Review in Wi-Lan matter.

Date	2013	InterDigital
To:	2015	
Duties:	Consulting expert in support of licensing a patent portfolio.	
Date	2013	Apple
Duties:	Consulting expert in National Cheng Kung University matter.	
Date	2013	Acer, ASUS, TPV
To:	2014	
Duties:	Testifying expert in Florida Atlantic University matter.	
Date	2013	Toshiba
To:	2014	
Duties:	Testifying expert in WiLAN matter.	
Date	2013	Grass Valley Group
Duties:	Consulting expert for defense in FastVDO matter.	
Date	2013	TI
To:	Present	
Duties:	Consulting expert.	
Date	2012	Vizio
Duties:	Consulting expert in Innovus Prime matter.	
Date	2012	Haier & Shandong Provincial Government
To:	2014	
Duties:	Foreign expert advisor for developing and acquiring patent portfolios for innovative consumer products. Development of a patent pool for patents essential to standards for wireless power.	
Date	2012	Vizio
To:	2014	
Duties:	Consulting expert in MIT matter.	
Date	2012	Nokia
To:	2013	
Duties:	Consulting expert.	
Date	2012	TPV & Vizio
To:	2013	
Duties:	Testifying expert for defense in Hitachi matter.	
Date	2012	Samsung
Duties:	Consulting expert in Myport matter.	

Date	2012	Verizon Wireless (Cellco) & T-Mobile
To:	2013	
Duties:		Testifying expert for defense in Realtime Data matter.
Date	2011	Apple
Duties:		Consulting expert in Technicolor matter.
Date	2012	Microsoft
To:	2014	
Duties:		Consulting expert for plaintiff in Motorola matter.
Date	2012	Apple & LG
To:	2014	
Duties:		Testifying expert for defense in MPT matter.
Date	2011	Sanyo & Olympus
Duties:		Consulting expert for defense in FastVDO matter.
Date	2011	Apple
Duties:		Consulting expert for defense in FastVDO matter.
Date	2011	Sony
Duties:		Consulting expert for defense in FastVDO matter.
Date	2011	Philips
To:	2014	
Duties:		Consulting expert for licensing essential patents.
Date	2011	Apple
Duties:		Consulting expert for defense in MobileMediaIdeas matter.
Date	2011	Video264Innovations
Duties:		Consulting expert.
Date	2011	Nikon, Fujifilm & Panasonic
Duties:		Consulting expert for defense in FastVDO matter.
Date	2011	Haier
To:	2014	
Duties:		Consultant for standards, patent pools and patent licensing.
Date	2011	Susman Godfrey
Duties:		Consultant for patent analysis.
Date	2011	Vizio

	Duties:	Testifying expert for defense in MPT matter.
Date	2011	Tivo
	Duties:	Consulting expert.
Date	2011	Nikon
To:	2013	
	Duties:	Consulting expert for defense in MPT matter.
Date	2011	Broadcom
	Duties:	Consulting expert.
Date	2010	Lionsgate Entertainment
	Duties:	Consulting expert in MPEG licensing matter.
Date	2010	Verizon
To:	2012	
	Duties:	Testifying expert for defense in ActiveVideo Networks matter.
Date	2010	Lenovo
To:	2011	
	Duties:	Consulting expert for defense in licensor matter.
Date	2010	Sonic, Cyberlink & Joint Defense Team
	Duties:	Consulting expert for defense in Mediostream licensor matter.
Date	2010	Wistron
To:	2011	
	Duties:	Consulting expert for defense in Toshiba licensor matter.
Date	2010	Panasonic, Pioneer
To:	2011	
	Duties:	Consulting expert for defense in VES licensor matter.
Date	2010	Zoran
To:	2011	
	Duties:	Technical analysis of patents potentially related to video coding standards.
Date	2010	Intravidual
To:	2015	
	Duties:	Consulting expert for patent licensing.
Date	2009	Apple
	Duties:	Technical analysis of patents potentially related to video coding standards.

Date	2009	SanDisk
Duties:		Consulting expert re Audio and Video coding patents
Date	2009	Lenovo
To:	2011	
Duties:		Consulting expert for defense in MPEGLA licensor matter.
Date	2009	Polycom
To:	2010	
Duties:		Consulting expert for defense in MPT matter.
Date	2009	DirecTV
To:	2011	
Duties:		Consulting expert for defense in MPT matter.
Date	2009	EchoStar
To:	2010	
Duties:		Consulting expert for defense in MPT matter.
Date	2009	Vizio
Duties:		Consulting expert for defense in Sony matter.
Date	2009	Vizio
To:	2010	
Duties:		Consulting expert for defense in LG matter.
Date	2009	Vizio, TPV, Amtran
To:	2010	
Duties:		Technical report & video presentation to US Customs on design issues related to Funai matter. Testified in ITC hearing, March 2010.
Date	2008	Humax (Korea)
To:	2009	
Duties:		Audio patent and prior art review.
Date	2008	IC Media (Canada)
To:	2009	
Duties:		Design of digital signage.
Date	2008	AVS Patent Pool Administration (China)
To:	Present	
Duties:		Co-Director, leading license negotiations among domestic and multinational licensors.

Date	2008	DirecTV
Duties:		Prior art search for DirecTV's appeal against Finisar.
Date	2008	Pure Digital
Duties:		Prior art research and analysis.
Date	2008	Wragge (UK)
Duties:		Prior art review for two video coding patents.
Date	2008	Microsoft
Duties:		Extensive non-infringement analysis for patents asserted by Multimedia Patent Trust (MPT).
Date	2008	iPotential
To:	2010	
Duties:		Technical analysis of patents potentially related to video coding standards.
Date	2008	Alba (UK) aka Harvard Research
To:	2009	
Duties:		General education on essential patents, licensing, prior art related to the MPEG2 video standard.
Date	2008	Philips (Holland)
Duties:		Review of patents for essentiality.
Date	2008	Proview
Duties:		Prior art search for TV display features.
Date	2008	Nissim
To:	2014	
Duties:		Consultant expert for plaintiff in Time-Warner case concerning patents related to DVD navigation.
Date	2008	Comcast
Duties:		Technical analysis of MPEG2 systems, DSM-CC, and carousel operation in the cable TV network environment. Contribution to expert's report in Comcast's successful litigation against Finisar.
Date	2007	Sun Microsystems
To:	2009	
Duties:		Chief architect of royalty-free video codec: OMS Video. Extensive supporting patent analysis and technical publication analysis for prior art review.
Date	2008	SanDisk

	Duties:	Technical analysis of patents and prior art searching for audio coding.
Date	2008	Stamberger (Munich)
	Duties:	Technical analysis of patents potentially related to video coding standards.
Date	2008	Trigem America Liquidating Trust
	Duties:	Preparation of expert reports concerning litigation between Toshiba and Trigem.
Date	2007	Susman-Godfrey
	Duties:	Technical analysis of patents potentially related to video coding standards.
Date:	2007	Wong-Cabello
	Duties:	Technical analysis of patents potentially related to video coding standards.
Date:	2007	Sisvel
	Duties:	Technical assistance in licensing audio patents to a third party.
Date:	2007	Quanta
	Duties:	Technical assistance in litigation related to DVD.
Date:	2006	CCTV (China Central TV)
	Duties:	Technical analysis of patents potentially related to MPEG2, H.264, AVS and VC-1 video coding standards and MPEG-AAC and Dolby audio coding standards.
Date:	2006	MediaTek
	Duties:	Technical analysis of MediaTek patents potentially related to video coding standards.
Date:	2006	Alcatel
	Duties:	Technical analysis of Alcatel patents potentially related to video coding standards.
Date:	2006	Motorola
	Duties:	Analysis of technologies contributed to the H.264 standard.
Date:	2005	Nokia
	Duties:	Analysis of technologies contributed to the H.264 standard.
Date:	2005	Avid
	Duties:	Technical analysis of patents potentially related to coding standards.

Date:	2005	IBM
Duties:		Analysis of technologies contributed to the H.264 standard.
Date:	2005	HP
Duties:		Analysis of selected audio coding patents.
Date:	2005	ComVentures
Duties:		Due diligence review of a start-up company in the video semiconductor business.
Date:	2005	nVidia
Duties:		Prior art search for a set of third-party patents.
Date:	2004	Microsoft
Duties:		- Working with a major financial services company, developed a procurement model for managing licensing costs. - Assisting development of non-infringement positions for a set of patents. (Not in suit).
From:	2004	SAGEM (France)
To:	2005	
Duties:		Technical analysis of potential infringement by SAGEM of a set of third-party patents.
From:	2004	Polycom Inc.
To:	2005	
Duties:		Technical analysis of Polycom patents potentially related to video coding standards.
From:	2004	Sony Corp.
To:	2006	
Duties:		- Technical analysis of potential infringement by Sony of a set of third-party patents. The parties settled without litigation. - Analysis of a third-party patent per Sony's direction.
Date:	2004	On2 Inc
Duties:		Delivery of the comprehensive set of reports on IPR potentially related to the emerging H.264 standard.
Date:	2004	Matsushita Avionics
Duties:		Technical analysis of potential infringement of Matsushita patents by a third party.
Date:	2004	Connex Inc.
Duties:		Teaching the details of key complex algorithms in the H.264 standard.

Date:	2004	DoCoMo (Japan)
Duties:		Delivery of the comprehensive set of reports on IPR potentially related to the emerging H.264 standard.
From:	2004	DivX Inc.
To:	2006	
Duties:		Analysis of technologies contributed to the H.264 standard
Date:	2004	Cues, Inc
Duties:		Assisted evaluation of infringement by Cues' products of a patent asserted by a third party. The parties settled without litigation.
From:	2003	GE Licensing
To:	2004	
Duties:		Working with a major financial services company, developed a business, market and financial model for valuation of a portfolio of patents through their expiration.
From:	2003	Institute of Computing Technology, Chinese Academy of Sciences
To:	Present	
Duties:		• Preparation of IPR Report for IPR potentially related to the AVS Video standard. This effort involved review of over 2500 abstracts and over 450 patents. • IPR Subgroup Chairman of AVS. Formed an IPR Experts group to establish a legal foundation for AVS IPR. This group comprised approximately 20 attorneys and business development managers from eight multinational companies representing the PC, consumer, semiconductor and communication industries. (Broadcom, IBM, Intel, Matsushita, Microsoft, Nokia, Sony, Sun Microsystems). The group drafted a member agreement, bylaws and a formal IPR policy governing licensing commitments, disclosure obligations, and protection of license holder rights. AVS Members approved these documents in September 2004. • Co-Director of AVS Patent Pool. Leading the drafting and negotiation of patent pool management and licensing agreements. • Adjunct Professor, supervising a graduate student project on deinterlacing algorithm development.
Date:	2003	Yahoo
Duties:		Prior art search to provide background information per Yahoo's request.
Date:	2002	Gateway, Inc.
To:	2007	
Duties:		Extensive prior art search and assistance with claim construction in a litigation matter with Lucent Technologies.

From:	2002	Apple Computer
To:	2007	
Duties:		• Technical analysis of Apple patents potentially related to H.264. • Prior art search for various patents per Apple's request.
Date:	2002	Siemens (Germany)
Duties:		• Technical analysis of Siemens patents potentially related to H.264. • Delivery of the comprehensive set of reports on IPR potentially related to the emerging H.264 standard.
Date:	2002	JVT Committee (Apple, Cisco, Nokia, Polycom, Sun Microsystems, Tandberg, Teles, Texas Instruments, UBVideo, VideoLocus)
Duties:		Developed a comprehensive set of reports on IPR potentially related to the emerging H.264 standard. The reports included a technical review of legacy and new patents for video coding, and analysis of the patent statements provided by all companies whose contributions were adopted into H.264.
Date	2002	Intervideo
Duties:		Advised on market opportunities in the emerging MPEG4 marketplace.
Date:	2001	Proskauer Rose LLP
Duties:		Provided background support in MPEGLA's (successful) effort to license MPEG2 to the major PC manufacturers.
Date	2001	Pixonics
Duties:		Developed marketing plan and marketing support materials for this incubator-stage company.
Date	2001	Eltron, (Taiwan)
Duties:		Provided overview of the digital video semiconductor market.
Date	1999	nDSP Corporation
Duties:		Drafted the sales and marketing plan for this early-stage company. Joined nDSP full-time.
Date:	1998	IBM Corp.
Duties:		Assisted with participation in the ATSC DASE (datacasting) standard.
Date	1998	Spruce Technologies
Duties:		Assisted with transition of Cagent MPEG encoder product after it was sold to Spruce.
From:	1997	Samsung Advanced Institute of Technology
To:	1999	

Duties: Supported Samsung researchers in MPEG4 standard development. Led effort to provide error-resilient 3D-mesh coding.

Date 1997 **MPEGLA LLC**
Duties: Review of opportunity for an MPEG4 patent pool

From: 1993 **CableLabs**
To: 1994
Duties: Teamed with a patent attorney to develop the initial essential patent list for MPEG2. Reviewed approximately 10,000 abstracts and 1,000 patents for the MPEG2 Video and MPEG2 Systems standards. This effort led to the MPEGLA patent pool.

Litigation Support Experience

Quinn Emanuel Urquhart & Sullivan, LLP, San Francisco, CA
Case Motorola Solutions vs. Verkada, Investigation No. 337-TA-1281
Project: Expert report, deposition and trial testimony on invalidity of Motorola patents.
Status: Final determination found Verkada not infringed
Filed: August, 2021
Location: United States International Trade Commission, Washington DC

Orrick, Herrington & Sutcliffe, LLP, New York, NY
Case ClearPlay vs. Dish Network & EchoStar Technologies, Case No. : 2:14-cv-00191
Project: Expert report, deposition and trial testimony on invalidity of ClearPlay patents.
Status: ClearPlay patents not infringed
Filed: March 13, 2014
Location: United States District Court for the District of Utah

Akin Gump, Fish & Richardson, Finnegan, Baker Hostetler, Washington, DC
Case Broadcom vs. Vizio, LG, Funai, MediaTek, MStar, Investigation No. 337-TA-1047
Project: Expert report and deposition on non-infringement of Broadcom's patent.
Status: Final Determination in favor of Respondants.

Jones Day, Cleveland, OH
Case Avid Technology vs. Harmonic, Case No. C.A. No. 12-627-GMS
Project: Expert reports and deposition on infringement of Avid's patent.

Status: Parties settled before trial.

Haynes and Boone, Dallas, TX
Case Realtime vs. Riverbed, Case No. 6:1S-CV-463-RWS-JDL
Project: Expert report, deposition and trial testimony (E. District Texas) on
invalidity of Realtime patent.
Status: Jury found Realtime patent invalid at trial.

Akin Gump, LLP, Irvine, CA
Case Personalized Media Communications “PMC” vs. TPV et al, Case No.
2:15-cv-01366-JRG-RSP
Project: Declaration in support of claim construction brief.
Status: Parties settled.

Feinberg, Day, LLP, Menlo Park, CA
Case Parthenon Unified Memory Architecture “PUMA” vs. Apple, Case No.
9:12-cv-80701
Project: Expert report and deposition concerning unified memory architecture.
Status: Trial stayed at request of PUMA. Parties settled.

Carey, Rodriguez, LLP, Miami
Case *Twentieth Century Fox Home Entertainment, LLC, Paramount Pictures
Corp., vs. Nissim Corp., Case No. 14-81349/813450-cv-MARRA
Project: Expert witness for patent owner in counterclaims for infringement.
Declaration and expert testimony at Markman Hearings concerning DVD
standards essential patents.
Status: Parties settled before trial.

*Nissim asserted its patents against Twentieth Century Fox and
Paramount. The latter filed a complaint seeking a declaratory judgment
of non-infringement and invalidity.

**TechKnowledge Law Group, Turner Boyd LLP & O’Melveny &
Myers LLP, Los Angeles,**
Case Florida Atlantic University Research Corporation vs. Acer, ASUS &
TPV, Case No. 9:12-cv-80701
Project: Expert testimony at Markman Hearings concerning scan converter
design. Expert reports and deposition testimony.
Status: Parties settled before trial.

Finnegan, Henderson, Farabow, Garrett & Dunner, LLP
Case WiLAN vs. Toshiba, Case No. 1:12-Civ-23744
Project: Expert reports & deposition testimony concerning digital television
design.
Status: Parties settled before trial.

O’Melveny & Myers LLP, Los Angeles & Akin Gump LLP, Dallas

Case Hitachi vs. TPV, Case No. 2:10-CV-260-JRG
Project: Expert reports, depositions & trial testimony (E. District Texas) concerning digital television design.
Status: Jury found no infringement of Hitachi's patents by TPV and found Hitachi's patents invalid. (Partially reversed by JMOL)

Fish & Richardson P.C. San Diego
Case Multimedia Patent Trust vs. Apple & LG, Case No. 10-cv-2618 H
Project: Expert report, & deposition testimony concerning development of technology, standards and patents for video coding.
Status: Jury found no infringement of MPT's patents and found MPT's patents invalid.

Gibson Dunn LLP, Palo Alto
Case Realtime Data vs. Verizon Wireless & T-Mobile, Case No. 6:10-cv-00493-LED
Project: Expert reports, depositions and testifying expert in jury trial (E. District Texas) concerning infringement and validity of patents related to data transmission.
Status: Jury found no infringement of Realtime Data's patents by T-Mobile, and found Realtime Data's patents invalid.

O'Melveny & Myers LLP, San Francisco
Case MobileMediaIdeas vs. Apple, Case No. 10-258-SLR
Project: Expert reports and deposition, concerning infringement and validity of patents related to MPEG Audio Layer III ("mp3") coding.
Status: Trial bifurcated.

Jones Day, Cleveland
Case Multimedia Patent Trust Vs. Vizio, Case No. 09-CV-0278-H (CAB)
Project: Expert report, concerning infringement of patents related to H.264 video coding. (Slated for deposition and jury trial)
Status: Parties settled on the day before my deposition.

Kellog, Huber, et al, Washington, DC
Case Active Video Vs. Verizon, No. 2:10-cv-00248-RAJ-FBS; 2011
Project: Expert report, deposition and testifying expert in jury trial concerning infringement of patents related to video-on-demand
Status: Jury finding of infringement with significantly reduced damages and no finding of willful infringement.

Jones Day, Washington, DC
Case Funai Vs. Vizio, et al, ITC Investigation No. 337-TA-617; 2010
Project: Expert report, deposition and testifying expert in ITC hearing concerning infringement of a patent related to ATSC broadcast standards
Status: Parties settled after the hearing following reports adverse to complainant.

Wilson Sonsini, Goodrich and Rosati, Palo Alto, CA.

Case Terayon vs. RGB Networks; 2007
Project: Expert report for defendant in a contract dispute alleging violation of a non-compete agreement.
Status: Mediation hearing resolved the dispute to the satisfaction of both parties.

Wilmer Cutler Pickering Hale and Dorr, Boston, MA

Case Qualcomm Inc. vs. Broadcom Corp. No. 05 CV 1958 B (BLM); 2007
Project: Expert report and deposition on behalf of Broadcom against assertions of infringement of patents concerning video encoding and decoding, and on Qualcomm's compliance with the policies and procedures for IPR policy in the H.264 standards committee (JVT of ITU-ISO/IEC).
Status: Broadcom found not guilty of infringing Qualcomm's patents. The patents were determined to be not essential for implementing the H.264 standard. Qualcomm found guilty of inequitable behavior in the manner of their participation in the H.264 standards committee. Qualcomm's patents held unenforceable against H.264.

Greenberg Glusker Fields Claman Machtinger & Kinsella.

Case DTS vs. MediaTek; 2005
Project: Expert report for defendant in a contract dispute related to hacking into semiconductor chips to enable functionality for which a license fee was payable.
Status: Mediation hearing resolved the dispute to the satisfaction of both parties. The Mediator personally thanked me for the clarity of my report, which covered technical, business and licensing issues.

Arvay Finley Barristers, Victoria, BC, Canada.

Case AVT Audio Visual Telecommunications vs. University of British Columbia; 2005
Project: Expert report and testimony on behalf of AVT in a contract dispute related to mobile video communication including failure to disclose intellectual property.
Status: Arbitration hearing resulted in adverse ruling to AVT.

Thelen Reid & Priest, San Jose, CA.

Case Toshiba Corp. vs. Trigem Computer Inc. & Trigem America Corp. Case No. CV-03-4558-GHK(AJWx); 2005
Project: Expert report on behalf of Trigem regarding invalidity of patents relating to video and sub-picture data in multimedia computer systems.
Status: Suspended due to insolvency of Trigem.

Morrison & Foerster LLP, San Francisco, CA.

Case	<u>Tivo Inc. vs. Echostar Communications Corp., Echostar DBS Corp., Echostar Technologies Corp., & EchoStar LLC, Case No. 2-04CV-01; 2005</u>
Project:	Expert report and deposition on behalf of Echostar against assertions of infringement of a patent concerning personal video recorders (PVR).
Status:	Adverse ruling to Echostar.
	Pinsent Curtis Biddle, Leeds, UK
Case	<u>Thomson Licensing S.A. and Deutsche Thomson Brandt GmbH vs. Pace Micro Technology PLC. Case No. HC 2000 No. 1584; 2001</u>
Project:	Expert report and slated to testify on behalf of Pace arguing invalidity of patent and defending against assertions infringement of a patent concerning interlaced video coding.
Status:	Parties agreed to settlement on the second day of trial. Settlement terms were very favorable to defendant, and shortly after the trial, Thomson joined the MPEGLA patent pool.

Education

1974	University of Sussex, England	Ph.D., Applied Sciences (EE/CS) Thesis on “Orthogonal Transform Coding of Still and Moving Pictures”. All research performed in residence at the Image Processing Institute, USC Los Angeles, CA.
1970	University of Liverpool, England	Bachelor of Engineering with Honours, Physical Electronics

Publications (Partial List)

C Reader, “Orthogonal Transform Coding of Still and Moving Pictures”, Ph.D. thesis, U. Sussex, England, 1973.

C. Reader, “Intraframe and Interframe Adaptive Transform Coding.” SPIE Vol. 66, pp. 108-117, *Efficient Transmission of Pictorial Information* (1975).

C. Reader, W.H. Chen, “Local Retransmission of Compressed Imagery Data”, SPIE Vol. 149 pp. 196-204, *Applications of Digital Image Processing* (1978).

L. Hubble, C. Reader, “State of the Art in Image Display Systems”, SPIE Vol. 199, pp 2-8, *Advances in Display Technology* (1979).

C. Reader, W. D. Flanagan, “Design Considerations for Real-Time Image Processing”, SPIE Vol. 180, *Real-Time Signal Processing II*, (1979).

C. Reader, L. Hubble, “Trends in Image Display Systems”, Proc. IEEE, Vol. 69, No. 5, pp. 606-614, (1981).

J. Adams, C. Patton, C. Reader, D. Zamora, “Hardware for Geometric Warping”, Electronic Imaging, April 1984.

Digital Image Processing Techniques (Computational Techniques), Michael P. Ekstrom (Editor), Chapter describing hardware for real-time image processing and display systems, Academic Press, 1984, pp. 289-360.

C. Reader, et al, “A New Architecture for Real-Time Image Processing,” SPIE Vol. 534, pp. 72-78, *Architectures and Algorithms for Digital Image Processing II* (1985).

B. Thacker, R. Rinaldi, C. Reader, “Multiwindow Displays of Text, Graphics and Images”, SPIE Vol. 757, pp. 104-110, *Methods of Handling and Processing Imagery* (1987).

Keynote opening speech at IEEE Conference on Image Processing, 1994.

C. Reader, “MPEG4: Coding for Content, Interactivity, and Universal Accessibility,” Optical Engineering, Vol. 35, No. 1 pp. 104-108 (1996).

C. Reader, P. Schirling, A. Puri, MPEG Tutorial paper at IEC – NCF/Infovision '97.

O. Avaro, P.A. Chou, A. Eleftheriadis, C. Herpel, C. Reader, J. Signes, “The MPEG4 Systems and Description Languages: a Way Ahead in Audio Visual Information Representation”, Signal Processing: Image Communication, Vol. 9, No. 4, pp. 385-431 (1997).

C Reader, “What MPEG-4 means to me”, Sig. Proc.: Image Comm. 9(4): 467, 1997

MPEG Video Compression Standard, Edited by Mitchell, Pennebaker, Fogg, Le Gall. Chapman & Hall, 1997. Chapter describing the process of establishing the MPEGLA patent pool, Chapter 16, *MPEG Patents*, pp.357-362.

W Gao, C Reader, F Wu, Y He, L Yu, H Lu, S Yang, T Huang, X Pan, “AVS - The Chinese Next-Generation Video Coding Standard”, Proc. NAB Conf., 1994.

C Reader, “Origins of the performance of H.264/AVC: an account of the development of H.263 and H.264 coding tools”, SPIE Vol. 5909, pp59090W-1-8 Applications of Digital Image Processing XXVIII (2005).

C Reader, “AVS Intellectual Property Rights (IPR) Policy, J. Comput. Sci. & Technol., Co. 21, No. 3 (May 2006).

Q Huang, W Gao, D Zhao, C Reader, “A Real-Time Video Deinterlacing Scheme for MPEG-2 to AVS Transcoding”, Advances in Multimedia Information Processing - PCM 2006, Vol. 4261/2006, Springer Berlin / Heidelberg.

W Gao, S Ma, C Reader, “An End-to-End Application System of AVS: AVS-IPTV in China”. PCM 2007: 50-57

C Reader, “Technical design & IPR analysis for royalty-free video codecs”. PCS 2010: 102-105

J Chen, F Xu, Y He, J D Villasenor, Y Han, Y Xu, Y Rong, C Reader, J Wen, “Efficient Video Coding Using Legacy Algorithmic Approaches”. IEEE Trans. Multimedia 14(1): 111-120 (2012)

W Gao, T Huang, C Reader, W Dou, X Chen, "IEEE Standards for Advanced Audio and Video Coding in Emerging Applications", Computer, vol.47, no. 5, pp. 81-83, May 2014.

Ma S, Huang T, Reader C, Gao W, AVS2-Making video coding smarter [Standards in a Nutshell], IEEE Signal Processing Magazine, vol. 32, no. 2, pp. 172 - 182, 2015.

Wang R, Huang T, Park S-H, Kim J-G, Jang E S, Reader C, Gao W, The MPEG Internet Video-Coding Standard [Standards in a Nutshell], IEEE Signal Processing Magazine, Vol. 33, No. 5, pp. 164-172, Sep. 2016.

Reader C, Patent Landscape for Royalty-Free Video Coding, SPIE Vol. 9971, Sep., 2016

Patents

<u>Patent Number</u>	<u>Date Issued</u>	<u>Title</u>
6,192,073	2000	Methods and apparatus for processing video data
5,845,112	2000	Method for performing dead-zone quantization in a single processor instruction.

Academic Positions

Taught “early bird” classes at Santa Clara University in Image Processing and Digital Signal Processing

Adjunct Professor of the National Engineering Laboratory for Video Technology, in the Computer Science Department at Peking University, Beijing, China.

Professional Associations and Achievements

- ISO/IEC Certificate of appreciation for development of the MPEG4 standard.
 - Senior Member IEEE.
 - Member SPIE.
 - Member ACM.
 - Member AES.
 - Member SMPTE.
-

Contact Information

- 1654 Hillstone Ave.
Escondido
CA 92029
- 408 867 4884 Mobile-Office
- cliffreader Skype
- CliffReader WeChat
- cliff@reader.com
- www.reader.com