

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

INTEL CORPORATION
Petitioner

v.

INTERDIGITAL, INC.
Patent Owner

IPR2024-01441
U.S. Patent No. 9,674,556

DECLARATION OF DR. CLIFFORD READER

Declaration

I declare that all statements made herein 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: _____



Clifford Reader, Ph.D.

CONTENTS

I.	Professional Background	3
A.	Summary	3
B.	Education	4
C.	Work Experience	4
D.	Standardization	6
E.	Intellectual Property Rights	7
F.	Curriculum Vitae	8
II.	Video Encoding and Decoding Technologies.....	9
A.	Analog Video	9
B.	Digital Video	10
C.	Video Compression	13
D.	Loop Filtering	16
1.	Introduction	16
2.	Video Frame Noise	23
3.	Video Frame (Image) Filters	25
4.	Loop filters	29
5.	Sparse Adaptive Non-linear Denoising Filters	31
6.	History of Loop Filtering in Video Coding Standards	34
III.	Relevant Legal Standards	37
IV.	Person of Ordinary Skill in the Art	39
V.	Overview of the Subject Patent US 9,674,556, (“’556 Patent”)	40
A.	Overview	40
1.	Field of Art	40
2.	Background - Problem statement, Admitted Prior Art	40

3.	Summary of the Alleged '556 Invention	44
4.	Detailed Description	44
B.	Claims	47
VI.	Prosecution History of the '556 Patent	50
VII.	Summary of Video Encoding and Decoding Prior Art.....	53
A.	U.S. Patent Publication No. 2003/0219073 (“Lee”).....	53
B.	Applicant Admitted Prior Art (“AAPA”): Adaptive sparse de-noising filter.....	54
C.	Applicant Admitted Prior Art (“AAPA”): Standard deblocking filter	55
D.	U.S. Patent Publication No. 2006/0110062 (“Chiang”)	56
E.	U.S. Patent Publication No. 2003/0152146 (“Lin”)	57
F.	SMPTE Standard: VC-1 Compressed Video Bitstream Format and Decoding Process, SMPTE 421M-2006 (“VC1 Standard”).....	57
G.	Technical Article: “A Nonlinear Loop Filter for Quantization Noise Removal in Hybrid Video Compression” (“Guleryuz”)	58
VIII.	Claim Construction	58
IX.	Ground IA – Claims 1–8 are Obvious under 35 U.S.C. § 103 by Lee	65
A.	Claim 1	65
B.	Claim 2.....	74
C.	Claim 3.....	76
D.	Claim 4.....	77
E.	Claim 5.....	78
F.	Claim 6.....	80
G.	Claim 7.....	80
H.	Claim 8.....	80
X.	Ground IB – Claims 1–8 Are Obvious under 35 U.S.C. § 103 Based on Lee in View of AAPA	81

A.	Claim 1	82
B.	Claim 2	87
C.	Claim 3	87
D.	Claim 4	88
E.	Claim 5	89
F.	Claim 6	91
G.	Claim 7	91
H.	Claim 8	91
XI.	Ground II – Claims 1–8 Are Obvious under 35 U.S.C. § 103 Based on Chiang in View of Lin	92
A.	Claim 1	92
B.	Claim 2	100
C.	Claim 3	102
D.	Claim 4	103
E.	Claim 5	103
F.	Claim 6	105
G.	Claim 7	105
H.	Claim 8	106
XII.	Ground III – Claims 1–8 Are Obvious under 35 U.S.C. § 103 Based on VC1 Standard in View of Guleryuz	106
A.	Claim 1	106
B.	Claim 2	116
C.	Claim 3	118
D.	Claim 4	119
E.	Claim 5	120
F.	Claim 6	121

G.	Claim 7	121
H.	Claim 8.....	122
XIII.	Summary	122
XIV.	Additional Remarks	122

I, Dr. Clifford Reader, do hereby declare:

1. I am making this declaration at the request of Intel Corporation in the matter of *Intel Corporation v. InterDigital, Inc.*, IPR2024-00574.

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:

- (a) U.S. Patent No. 9,674,556 (“’556 Patent”) (EX1001)
- (b) Prosecution History of the ’556 Patent (EX1002)
- (c) Prosecution History of U.S. Patent Application Publication No. 12/312,386, filed as PCT/US2007/22795 (“’386 Application”) (EX1005)
- (d) U.S. Patent Appl. Publ. No. 2003/0219073 (“Lee”) (EX1006)
- (e) U.S. Patent Appl. Publ. No. 2006/0110062 (“Chiang”) (EX1007)
- (f) U.S. Patent Publication No. 2003/0152146 (“Lin”) (EX1008)
- (g) SMPTE Standard: VC-1 Compressed Video Bitstream Format and Decoding Process, SMPTE 421M-2006 (“VC1 Standard”) (EX1009)

- (h) Onur G. Guleryuz, A Nonlinear Loop Filter for Quantization Noise Removal in Hybrid Video Compression, IEEE Int'l Conference on Image Processing, 2005 (EX1010)
- (i) U.S. Patent No. 8,194,757 to Schoner (EX1011)
- (j) U.S. Patent No. 6,360,024 to Tan (EX1012)
- (k) U.S. Patent No. 1,1796,030 to Kell (EX1013)
- (l) Xin Li & Michael T. Orchard, Spatially adaptive image denoising under overcomplete expansion, Proc. ICIP, Vol. 3, Sep. 2000 (EX1014)
- (m) ISO/IEC/JTC1/SC2/WG8 Moving Picture Expert Group, MPEG 89/254, Adaptive filter in the coding loop, Jan. 1990 (EX1015)
- (n) CCITT SGXV Working Party XV/1 Specialists Group on Coding for Visual Telephony, #356, Loop Filter Improvement, Sept. 19, 1988 (EX1016)
- (o) '556 Patent Infringement Contention Claim Chart, Exhibit 18 to InterDigital's First Suppl. Responses to Respondents' First Set of Interrogatories in ITC-337-TA-1373 (EX1024)
- (p) Commission Investigative Staff's Markman Brief, ITC-337-TA-1373 (EX1025)
- (q) U.S. Patent No. 7,388,999 ("Samadani") (EX1026)

- (r) ISO/IEC/JTC1/SG2/WG8 Moving Picture Expert Group MPEG 89/248 (EX1027)
- (s) Reader C, Intraframe and Interframe Adaptive Transform Coding, SPIE Vol. 66, 1975 (EX1028)
- (t) Reader C, MPEG4- Coding for Content, Interactivity, and Universal Accessibility (EX1029)
- (u) MPEG Video Compression Standard, Edited by J. L. Mitchell et al, Chapter 16 (EX1030)
- (v) '556 Patent Infringement Contention Claim Chart, Exhibit 5 to InterDigital's Infringement Contentions in *InterDigital et al. v. Lenovo Group Ltd. et al.*, no. 5:23-cv-00493-FL (E.D.N.C.) (EX1031)
- (w) Respondents' Markman Brief, ITC-337-TA-1373 (EX1032)
- (x) Corrected Complainant's Markman Brief, ITC-337-TA-1373 (EX1033)

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 thesis¹ and summarized in an SPIE paper.² 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.

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.

¹ Reader C., "Orthogonal Transform Coding of Still and Moving Pictures," U. Sussex, England, 1973.

² Reader C., "Intraframe and Interframe Adaptive Transform Coding," SPIE Vol. 66, 1975.

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 MPEG.

13. I was Head of Delegation (HoD) to MPEG for the United States in 1991-1992, and 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.

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 language³. MPEG4 focused on audiovisual low bitrate coding, including low bitrate speech coding.

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

17. 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.

E. Intellectual Property Rights

18. 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.⁴

19. In 2002 I was hired by 10 companies to evaluate the standards essential patent environment for the nascent H.264 standard. I was an invited expert to the

³ Reader C, "MPEG4: Coding for Content, Interactivity, and Universal Accessibility", Optical Engineering, Vol. 35, No. 1, pp. 104-08, Jan. 1996.

⁴ "MPEG Video Compression Standard", Edited by J. L. Mitchell et al, Chapter 16, "MPEG Patents," pp. 357-62.

Joint Video Team (JVT) established by ISO and ITU to develop the H.264 standard (also denoted MPEG4 Pt. 10, AVC).

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

21. In 2013 I was hired to evaluate the standards essential patent environment for the emerging AV1 standard

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

23. I am the Co-Director of the China AVS Patent Pool Administration. I lead the negotiations for sub-licensing of the AVS standards.

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

F. Curriculum Vitae

25. 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 set forth 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. Video Encoding and Decoding Technologies

26. Below I describe the state of video encoding and decoding as of the '556 Patent's filing date and summarize the development of technologies underlying the features and methods recited in claims 1–8.

A. Analog Video

27. 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.

28. All-electronic television⁵ 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.

29. 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. Analog broadcasting continued in the U.S.

⁵ Electro-mechanical television systems were also introduced.

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

30. Analog video intrinsically contains noise from several sources. Television cameras introduce noise into the signal. If the production is on magnetic tape, the tape introduces further noise. This is exacerbated by any iterative post-production editing process.

B. Digital Video

31. Digitization of analog signals was introduced in the late 1940s with the name “pulse code modulation” (PCM). Essentially, 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.

32. 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.

33. 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.

34. 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 resolution 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 resolution to 1000 lines, and the recent so-called 4K systems double the resolution again to 2000 lines. The imminent next generation of so-called 8K systems will further double the resolution to 4000 lines, which will approximate the resolution of a 35mm film system.

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

36. 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.

37. 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.

38. 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.

39. Video coding 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.

C. Video Compression

40. One of the earliest examples of video compression used change detection to affect the removal of redundancy. For example, U.S. Patent No. 1,796,030 to Ray D. Kell, filed in 1929, describes the efficient transmission of video by only transmitting changed elemental areas of the image from one frame to the next. Kell's work was further developed at Bell Labs. in the 1960s by a process called conditional update that distinguishes stationary from non-stationary pixels, and transmits only the pixels that have changed. Related work by F.W. Mounts, in 1969, described using a threshold to determine significant changes between pixels, where the threshold identifies the number of pixels having significant change required in order to trigger transmission of the change. Mounts also acknowledges the threshold that applies to distinguishing noise in the video from a true change.

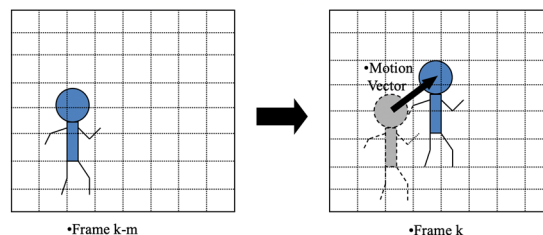
41. The image and video compression work at Bell Labs. was almost exclusively pixel-based in the 1960s and 1970s. The cost and practicality of memory was very high, meaning that buffering many rows of pixels to perform processing on two-dimensional blocks of pixels was prohibitive, and the cost of logic to process

blocks of pixels at real-time video rates was also prohibitive. Some researchers, however, pursued an approach called transform coding, in which blocks of pixels such as 16 x 16 pixels were processed to remove spatial and temporal redundancy. Instead of performing pixel-based processing, the transform approach transforms blocks of pixels into spatial frequency coefficients, which can be more easily and effectively coded than the raw pixels themselves.

42. Block-based motion compensation was developed in 1974. Modern video coding standards all use this technique, which is why semiconductor chips for video compression perform change detection.

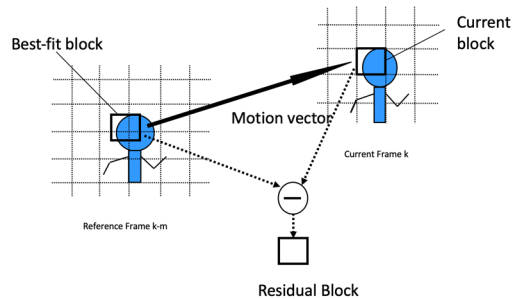
43. With block-based motion compensation, motion is modeled as a translation of objects within the scene, in the plane of the imaging system. While this model is clearly an approximation (for example, it ignores rotation, zoom and occlusion), it works remarkably well, especially when applied to small blocks of pixels into which each video frame may be divided.

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



44. The block-based motion compensation technique identifies blocks in the scene in which the content has not changed in the current frame from the previous

frame, and does not update them, while blocks that have changed are predicted from blocks in the previous frame that may be displaced. The displacement is called a motion vector.



45. Because the prediction may not be perfect, the pixel values of the predicting block are subtracted from the pixels of the block to be coded, yielding a residual block of prediction errors. If the simple translation model for motion is valid for the current block to be coded, the residual pixel values are zero or very small, and can be coded with few bits. It should be noted however that because the model can fail, a block of pixels may be coded more efficiently without reference to another frame, and a decision is made on a block-by-block basis whether to use temporal prediction – which is often called inter-prediction – or whether to predict the values of the pixels in the block from neighboring blocks of pixels in the same frame. Such spatial prediction is often called intra-prediction. It also may produce a block of residual pixel values. Modern video coding standards all use these techniques.

46. It should be noted that the use of prediction from one frame to the next means the encoder/transmitter and decoder/receiver must maintain identical copies

of the frames used for prediction. This implies the decoder must receive a perfect transmission of the differences between a current frame and a prediction frame, else it will not correctly reconstruct the current frame, and then the reconstructed current frame will not provide the correct reference to which differences for the ensuing frames will be added, meaning the error will propagate.

47. 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. Temporal redundancy is removed using block-based motion compensation. 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 the transform coding techniques, referred to earlier.

D. Loop Filtering

1. Introduction

48. For over 50 years, the principal approach to video coding has been based on block transformation of pixel data and quantization (i.e., a form of lossy compression that compresses a range of values to a single quantum (discrete) value) of transform coefficients. In simple terms, this is because natural video data contains a predominance of low spatial frequencies, and thus the transformation tends to concentrate the information spread across the entire spatial (pixel) domain into a

small number of coefficients in the transform domain, the information in which may then be efficiently coded.

49. Given that the correlation between pixels decreases rapidly with spatial distance, and that the implementation of transforms increases non-linearly with size, transform coding has been implemented by dividing each video frame into small blocks – e.g., 8x8 pixels – and performing the transformation and quantization operations on each block.

50. The goal of video coding is to represent the *information* in the video frames with as few bits as possible, for efficient transmission or storage, while providing for reproduction of the video frames with as minimal *perceptual* distortion as possible. In this context, germane to this matter at hand, a distinction is made between the *subjective* quality of the reproduced video data – i.e., what the human observer perceives – and the *objective* quality of the reproduced video data – i.e., a mathematical measure of the error between reproduced video pixel values and the original pixel values (e.g., mean squared error; “MSE”). So-called loop filters for video coding have focused on improving subjective video quality.

51. The perception of video quality is complex and may be counter intuitive. In video sequences of natural scenes, straight horizontal and vertical edges – e.g., the boundaries of blocks used for coding the video – will become visible (i.e., unnatural) if there is appreciable contrast between pixel values on either side of the

edges. Another example concerns the perception of errors if sharp edges are not reproduced accurately, and an original smoothly continuous edge is reproduced with visible pixels having incorrect values offset on either side of the edge. This is often called the “mosquito” effect. An illustration is provided in Fig. 2A of US8194757, reproduced below. Yet another example concerns coding of “flat” – smooth – areas of a video frame, in which if the pixel values are quantized too coarsely a contour effect similar to “painting by numbers” may be perceived.

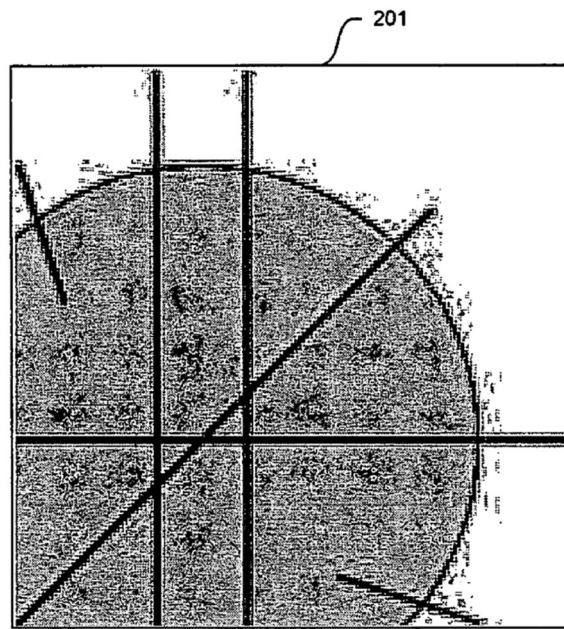


FIG. 2A

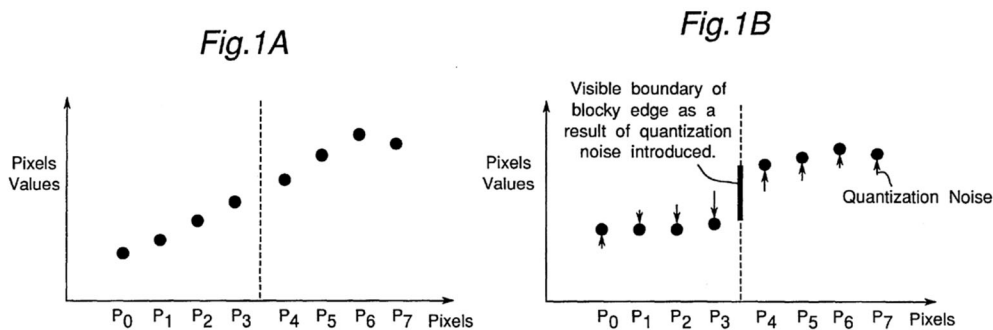
52. While video data comprises a succession of still-image frames, certain types of distortion that are imperceptible if each single frame is viewed as a static image, may become visible – even to the point of being highly objectionable – if the video frames are viewed as a real-time video sequence; and vice versa. For example,

small errors on reproducing edges in a single video frame, or reproducing slowly changing pixel values across a flat region of a single video frame, that are not perceivable, may become highly visible on reproduction of a succession of such frames in a video sequence, because the spatial location of the edge or the spatial location of the contours respectively may be changing from frame-to-frame, leading to a “busyness” or “jittering” of the perceived video. The classic example is the mosquito effect in which the “mosquitos” appear to be buzzing around the corrupted edge.

53. Nevertheless, video coding is performed by modeling video data mathematically – e.g., statistically – and using the measured statistics to control various modes and parameters of an adaptive coding process. This is based on the nature of the video data itself, which varies spatially across the extent of a video frame and temporally across the video frames comprising a video sequence. A key issue is the homogeneity of the video data, spatially and temporally, which in mathematical terms concerns the stationarity of the pixel data statistics. Considering the above-mentioned approach of coding video data in small blocks of pixels, this issue translates into the issue of whether pixel data is stationary within a pixel block and whether it is stationary between pixel blocks. To manage the inhomogeneous nature of video data, video coding has been performed adaptively on a block basis

since the early 1970s.⁶ This means some measure is made of the data locally and used to adjust parameters of a video coding algorithm, or select between alternate video coding algorithms.

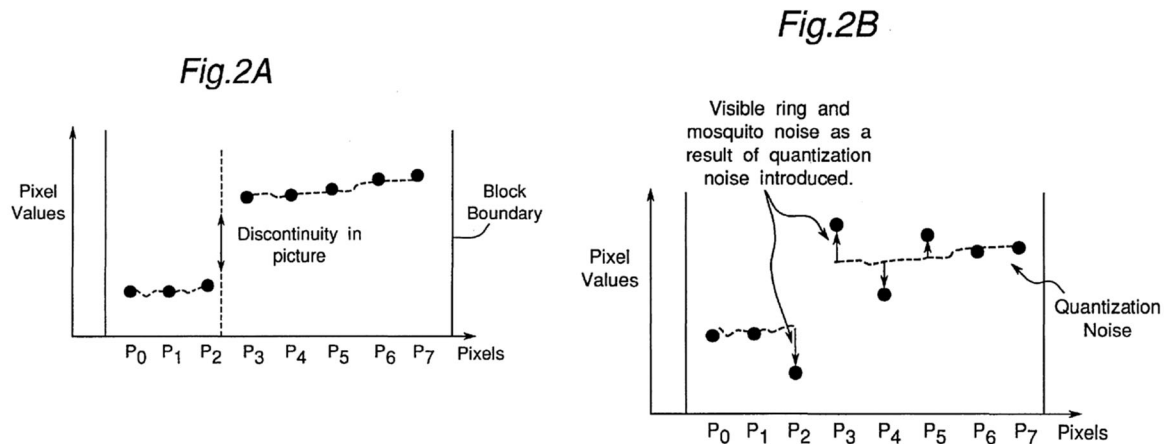
54. When blocks of video data are encoded, the question arises whether the continuous (smooth) change in pixel values that (usually) occurs in the original data across the boundary between two blocks is preserved when the respective blocks are reproduced. If it is not, and a sharp change in pixel values – e.g., brightness – occurs, the edge may become perceptible, particularly because it will be an unnatural horizontal or vertical straight line. An illustration is provided in Fig. 1 of US6360024, reproduced below. Of course, a natural edge that happens to be horizontal or vertical may happen to align with a block boundary, leading to the challenge of mitigating false edges while preserving natural edges. For this reason at least, adaptive filtering has been employed.



⁶ Reader C., Orthogonal Transform Coding of Still and Moving Pictures, Dec. 1973; summarized in Reader C, Intraframe and Interframe Adaptive Transform Coding, SPIE Vol. 66, 1975

55. A distinct issue is the homogeneity/stationarity of pixel data at large. A video frame typically comprises the projection of a three-dimensional world onto the two-dimensional focal plane of the imaging system. It comprises a set of objects superimposed on each other and perhaps a background. Each object may have distinct visual information characteristics, and corresponding statistical properties, thus coding the interior of each object may require adaptation to the local characteristics. In particular, the video data may comprise regions of “flat”/smooth pixel values, or uniformly “busy”/textured pixel values – both of which have reasonably homogeneous spatial information/statistics – but also contain sharp edges that separate homogenous regions and constitute inhomogeneous discontinuities/singularities.

56. Now considering the coding of such information using a regular array of blocks – the block boundaries may fall on a regular geometric grid, the boundaries of which have arbitrary location relative to the video information. The aforementioned sharp edges may fall within a block, at arbitrary orientations to the block grid. In this case, the block characteristics/statistics may be distinctly different on either side of the sharp edge. Fig. 2 of US6360024, reproduced below, illustrates this case, when the video coding does not code the sharp edge accurately.



57. It has been known since at least 1990 that these two problems are distinct, and that a combination of approaches in digital signal processing are needed to optimally address both.⁷ See extracts below.

Two kind of adaptive filters have been introduced before the frame memory.

FILTER ON THE BLOCK BOUNDARIES:

To avoid the blockiness and transitions on the block boundaries, due to the quantisation of the DC coefficients, a low pass filter has been introduced on the pixels around the 8*8 block boundaries.

x ₁	x ₂	x ₃		
			block edge	$d_1 = (x_1+x_2+x_3)/3 - (y_1+y_2+y_3)/3$
y ₁	y ₂	y ₃		
				$d_2 = (y_1+y_2+y_3)/3 - (z_1+z_2+z_3)/3$
z ₁	z ₂	z ₃		

If $|d_1| < T$ and $|d_2| < T$ the pixels x_2 and y_2 are filtered.

A (1/4 1/2 1/4) low pass filter is used vertically for horizontal edges and horizontally for vertical edges.

The threshold T can be chosen adaptively relative to the contents of the picture and to the step size.

⁷ Eindhoven, ISO/IEC/JTC1/SC2/WG8, Moving Picture Expert Group, MPEG 89/254, Adaptive filter in the coding loop, Jan. 1990 (EX1015).

FILTER OVER THE WHOLE FRAME:

To reduce the mosquito effects around the objects and the transitions, it is necessary to process not only the block boundaries, but also the entire pictures. For that, it is necessary to detect transitions in the frame. The idea is to compare the difference between neighbouring pixels and to decide if the difference is due to a transition or to noise. If the difference is less than a threshold, the pixels are filtered. The threshold has to be chosen relatively to the local step size. In a first approach, the value of $\text{step}/2$ has been chosen.

58. During the 1980s and 1990s, extensive research was conducted by many parties into these problems, including investigating “loop filtering” the data inside the video coding loop, i.e., after each frame was decoded and before it was stored in the reference frame buffer for use in predicting future coded frames, or “post filtering” the data after it was decoded and output for display, but not filtering the reference frame. As far back as 1986, it was recognized that filtering inside the loop resulted in reference frame data that produced more efficient motion estimation/compensation of future frames. Work throughout the 1990s repeatedly confirmed this result.

2. Video Frame Noise

59. In traditional video coding, a set of video coding tools is used to collectively address the problem of isolating the spatial information within the arrays of pixels forming each video frame and isolating the temporal information in the succession of frames forming a video sequence in order to code the information with the maximum efficiency. Typically, coding tools perform lossless coding except one

– quantization of the transform coefficients. The mathematically non-linear operation of quantization means it is not invertible – the encoder introduces errors the decoder cannot recover from. As discussed earlier, the art of video coding is to minimize the number of bits representing the information in the video sequence, while also minimizing the perceptual error on the reproduced video sequence. As also noted, the video coding tools operate on small blocks of video pixels, and the interaction between this block structure and the video coding tools can strongly impact the perceived errors. To the extent such errors inevitably arise, loop filtering has been used to mitigate the visibility of the errors.

60. One source of errors can be the use of finite-length transforms – i.e., transforms that span only the dimension of the small pixel blocks. Use of orthogonal transforms themselves does not introduce error – the full spectrum of coefficients can be inversely transformed to exactly reconstruct the original data – but when quantization coarsely represents at least some coefficients, or truncates them, the values of the inverse transformed data will be different from the original – notably at the ends of the data, i.e., the boundaries of the blocks. Considering two adjacent blocks, a discontinuity of values may result, as illustrated earlier.

61. This issue may be exacerbated when adjacent blocks are coded using different quantization parameters (QP) and when different coding modes are used for adjacent blocks.

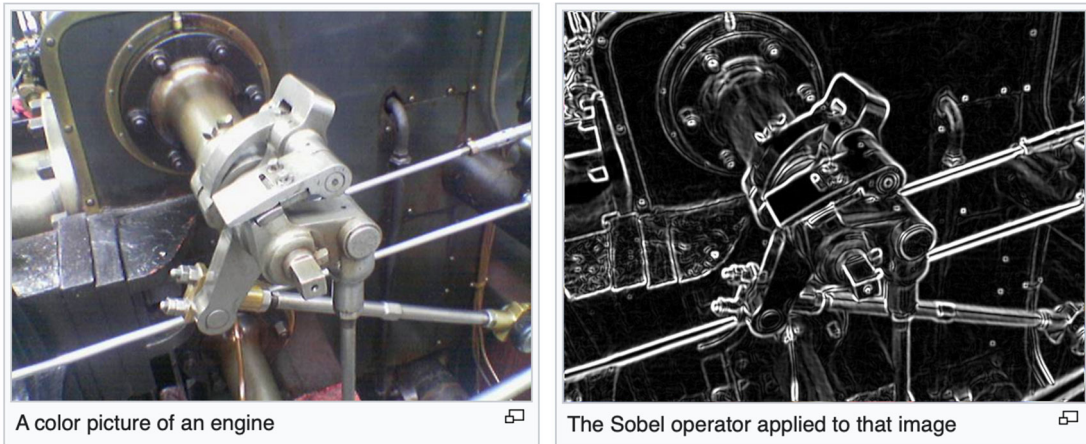
62. Across the entire video frame – notably in the interior of coding blocks – if the data are smooth (flat regions), the coding model will work well, and if the data are (even highly) textured, the coding model will work well, because although significant quantization error may be introduced, the human vision system is insensitive to errors in highly detailed areas. But the existence of sharp edges within the block will violate the model underlying the use of transform coding, expressed in the opening of this discussion above (that the statistics will be non-stationary). A sharp edge will result in a transform domain spectrum that contains significant energy/information in high-frequency coefficients, and if the non-linear quantization operator coarsely quantizes or truncates these coefficients, the edge will not be accurately reproduced. Two perceptually visible effects may result. The edge may not be as sharp – it may appear blurred/fuzzy; but also, due to the transformation, there may be “ringing”, i.e., a decaying oscillation of values on both sides of the edge. Fig. 2B of US6360024 above illustrates this effect. This may be perceived as the mosquito effect.

3. Video Frame (Image) Filters

63. Spatial filtering of video frames, following practices established for still images, has been performed for decades, using a variety of filter classes. A simple class is linear filters, implemented as a convolution of the pixel data with a linear filter kernel. Such kernels are typically two-dimensional square arrays of filter

coefficients, such as 3x3. The coefficients may be simple integer values that for example implement a spatial low-pass filter, with filter coefficient values chosen to normalize the values of the output pixels. Such filters may be non-adaptive, or the coefficients may be derived from local data and adapt as the filter moves across the data. Alternatively, the filter may be non-linear; for example, the filter may be conditionally operative depending on local conditions. A classic non-linear filter is the median filter which may be applied in one dimension on rows or columns of pixels or two-dimensionally on arrays of pixels.

64. A further example of non-linear filtering is the class of edge-detection filters, of which the Sobel filter (1983) is perhaps the best known. This 3x3 filter convolves a video frame to produce the edges of objects in the frame as an output. Because the distribution of edges in a frame is often a small percentage of the number of pixels in the frame, the data in the frame that actually produces an **output** is sparsely distributed, and the data of the output frame is “sparse.” But the filter itself is not sparse – it operates the same way on every input pixel of the original image, which itself is also not “sparse.” Illustrations of the Sobel filter operation can be seen on the Wikipedia page: https://en.wikipedia.org/wiki/Sobel_operator and are reproduced below.



65. This is distinct from a class of filters in which the *filter* itself is sparse, not the *data* that the filter operates on being sparse. The sparse filters can operate using sets of overcomplete transformations of the pixel data to produce sparse decompositions of the pixel data. The members of the set of transformations operate on displaced versions of the pixel data, which may be called a redundant representation. In the transform domain, various techniques have been used to non-linearly filter the transform coefficients, recognizing that the coefficients representing the uniformly distributed noise in the pixel domain are reasonably distinct from the coefficients representing the video frame data itself. In other words, multiple, slightly offset transforms of the image are taken and compiled. The coefficients of edges in the image data will have a high frequency and may have high magnitude, whereas the coefficients of the uniformly distributed noise will not. The filtering operation removes the low frequency coefficients, leaving only the image coefficients that are relatively “sparse” as compared to the uniformly distributed

noise coefficients. Following the filtering, the transform coefficients are inverse transformed and then the set of inverse transformed data are combined to produce an output frame that has substantially less noise. An excerpt from one of the papers⁸ is reproduced below to illustrate the process using an adaptive, sparse, denoising filter, in which it is the design of the filter which is sparse, not the data on which the filter is operating:

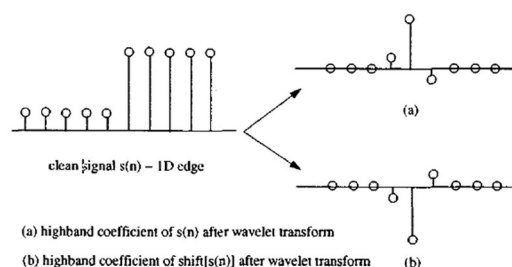


Fig. 1. Overcomplete expansion of a 1D step-edge signal.

Let us use a 1D step edge s , as an example to explain the advantage of using overcomplete expansion. Similar arguments also hold for 2D signal. Since wavelet transform is linear, the effect of noise on signal can still be treated as additive in wavelet domain. So we can separately study signal and noise components. On signal part, figure 1 shows the highband coefficients of all shifted version of s_n . Our interest actually only lies in large coefficients (spikes) since they carry the information about transitions. When step edge is shifted by one sample, the spike changes its sign but still remains as a spike. On noise part, one obvious fact is its correlated structure (e.g. version (a) and (b) are correlated). However, the noise is still independent of s , and its power is typically much lower than that of signal component (spikes). Consequently, the probability of both spikes disappearing is much less than [sic] the probability of either one disappearing, which explains why signal is better preserved under overcomplete expansion. Averaging over all

⁸ EX1014 (“Li”).

shifted versions is simply a convenient solution without resolving the uncertainty of in which version the spike is gone.

EX1014, p. 301. The Li, Orchard filter is an example of a sparse filter because the high frequency coefficient of the image data in the transform domain is “sparse” relative to the coefficient of the uniformly distributed noise, as depicted by the spikes illustrated in Fig. 1 of Li above.

66. It should be noted that the non-linear median filter mentioned above is a filter that may operate on sparse data. For example, if an image frame is corrupted by noise spikes on small percentage pixels, and thus sparsely and arbitrarily distributed throughout the frame, a median filter operating identically by convolution with every pixel in the frame may be designed such that only the sparse data of the noise spikes is filtered. This does not mean the median filter is a sparse filter.

4. Loop filters

67. Since the 1980s, spatial filtering has been used to improve the perceived video quality of reproduced video data. For example, during the development of the H.261 videoconferencing standard in the latter half of the 1980s, extensive study was undertaken. In brief, the decisions for incorporating loop filtering into that standard were conditioned by the complexity and cost of implementing filters into realtime encoding and decoding video terminals for conversation services with the semiconductor logic and memory available at that time. One proposal in 1988

suggested the use of two in-loop filters: a deblocking filter followed by a filter designed to remove random noise.⁹ Ultimately, a simple, linear filter was included in the H.261 standard, that can be enabled/disabled at the macroblock layer. (The committee necessarily performed a tradeoff between performance and implementation complexity/cost, given the requirement for realtime encoding and decoding for conversational services and the limited semiconductor technology of the time). The standards committee acknowledged the advantage of loop filtering for improving the efficiency of motion compensated prediction and chose to incorporate a loop filter, rather than a post filter.

68. Adaptive, non-linear deblocking filters were extensively studied in the 1990s. Designs involving one-dimensional (1D) filters and two-dimensional (2D) filters spanning horizontal and vertical block boundaries were developed and tested for use as loop filters and post filters. In brief, adaptive, non-linear, 1D loop filters were selected for superior performance and positive contribution to motion compensated prediction. These filters are adaptive in multiple ways:

- Decision whether or not to filter on a pixel basis along the boundary
- The length of the (1D) filter spanning the boundary
- The filter coefficient values
- The strength of the filter coefficient values

⁹ CCITT SGXV Working Party XV/1, Specialists Group on Coding for Visual Telephony, #356, Loop Filter Improvement, Sept. 19, 1988 (EX1016).

69. The determination of the adaptivity may be performed in multiple ways, for example:

- Differences in pixel values across the boundary
- Differences in pixel values on one side of the boundary compared to differences in pixel values on the other side of the boundary

70. The filter may be enabled/disabled at different levels of the video data hierarchy, such as the sequence level, picture level and sub-picture level. This allows for the application of the filter only at levels and in circumstances where the improvement to image quality is justified by the computational cost of the filtering.

71. In the latter half of the 1990s, attention also turned to the artifacts/noise inside video coding blocks, and deringing filters were studied for in-loop use and post processing use. The studies examined adaptive linear and non-linear, 1D and 2D filter designs.

5. Sparse Adaptive Non-linear Denoising Filters

72. A brief summary is provided of several references for the subject filters.

73. A 2000 paper by Li and Orchard¹⁰ summarizes prior work on denoising using the wavelet transform, but argues for the superiority of redundant representation using overcomplete expansions. In brief, the paper discloses use of shifted versions of the signal to provide the multiple, redundant representations.

¹⁰ EX1014.

The paper notes that “[i]mportant features in images (e.g. edges) are usually represented by large coefficients, thus facilitating their separation from noise components which usually correspond to small coefficients in wavelet domain” and further notes: “The highband coefficients are viewed as the mixture of non-edge coefficients observing a zero-mean model and edge coefficients observing a biased-mean model. The optimal noise suppression strategies(MMSE estimation) are derived for both edge and non-edge coefficients.” Li, 300.

74. In sections 3.1, and 3.2, Li discloses processing the distinct coefficients that represent the signal and the noise for non-edge coefficients and edge coefficients to achieve the goal of suppressing the noise while preserving the signal. Li, 302. For the edge coefficients, the paper also discloses a directional filtering approach to account for the orientation of the edge. *Id.*

75. A 2003 patent by Samadani¹¹ also employs shifter versions of the signal (images or video frames), in combination with transformations and non-linear filtering in the transform domain to denoise the signal data. Figures 3 and 4 provide an overview of the process.

¹¹ U.S. Patent No. 7,388,999 (EX1026) (“Samadani”).

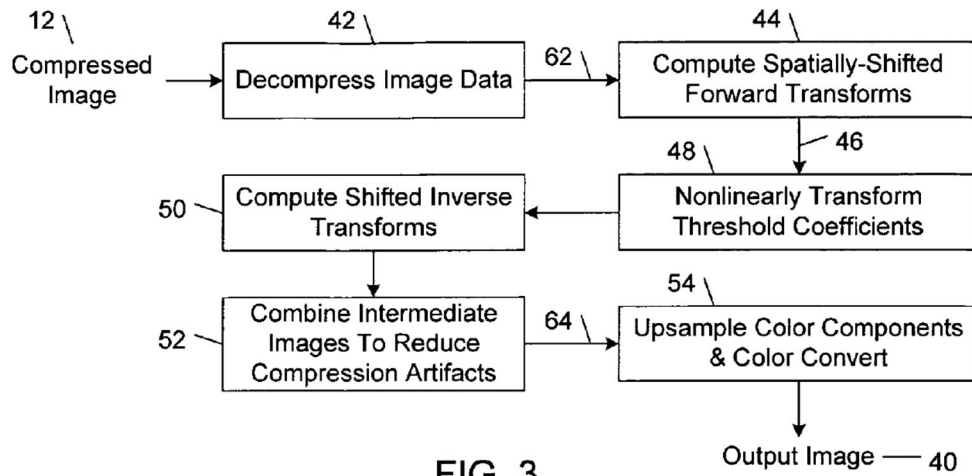


FIG. 3

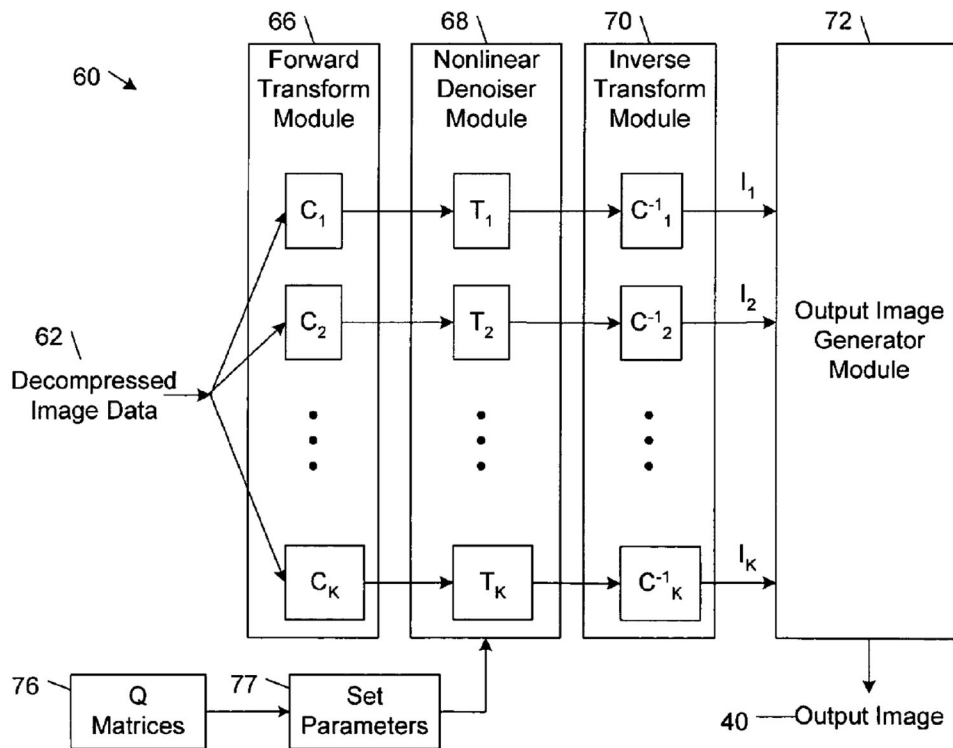


FIG. 4

76. The filtered transforms are inverse transformed and then combined. In the Samadani patent, the transform may be a discrete cosine transform (DCT). Non-linear thresholding and scaling operations may be performed on the transform

coefficients to perform the denoising function. Samadani, 7:8-8:30. The patent also discloses sharpening operations that may offset any blurring. Samadani, 8:31-63.

6. History of Loop Filtering in Video Coding Standards

77. As noted, the history of loop filtering in video coding systems goes back many decades. This section is a brief review of the timeline for this development, and the video coding standards incorporating loop filters.

78. During the development of H.261 in the 1980s, the committee developed a simple, linear 3x3 loop filter placed after the picture memory, to smooth the prediction data in Inter-coded macroblocks. The concept was to remove high-frequency artifacts caused by inaccurate motion vectors

79. The H.261 loop filter was included in many proposals for MPEG1 in 1989, but the MPEG committee subsequently decided to use motion compensation with ½-pixel accuracy and did not include the loop filter. The MPEG2 standard followed suit.

80. During the development of MPEG1, a proposal was made for an adaptive 1D filter with dependency on the quantization parameter (QP).¹²

81. Another proposal disclosed the use of two adaptive filters in the coding loop before the frame memory – one for deblocking and one for the “entire picture”

¹² EX1015.

for mosquito filtering.¹³ The deblocking filter is conditionally applied based on a threshold derived from pixel differences across the boundary, and upon QP. The whole frame mosquito filter adapts based on neighboring pixel differences and a threshold that depends on QP.

82. In the latter half of the 1990s, the ITU H.263 standard, and ISO MPEG4 Part 2 Visual standard were developed in parallel in a collaborative way. Both standards included deblocking filters; for example, the second version of H.263, published in 1998, included a simple adaptive 1D deblocking filter in optional Annex J.

83. The work in H.263 Annex J then evolved into the sophisticated deblocking filter in H.264 / MPEG4 Part 10 AVC, which was developed jointly by ITU VCEG (Video Coding Experts Group) and ISO MPEG. The first two versions of the H.264 standard were published in May 2003 and March 2005.

84. All of the video coding standards, beginning with H.261 in the 1980s, code video data in a hierarchical data structure. This is logical from the perspective that video is a multi-dimensional medium, that has to be stored or transmitted as a one-dimensional sequence, so there is a natural decomposition into an ordered sequence, since a complete item of content is a sequence of frames and the frames

¹³ EX1015.

comprise a sequence of blocks of pixels. Given that the video coding is highly adaptive, which means it is dependent on the video content itself, it follows that the choice of coding tool, the modes of such coding tools, and parameters for the coding tools may all vary based on the content, and such variation may be occurring at the block level, a region containing multiple blocks level, the frame level, or a sequence of frames level. The video encoder has access to the original video data and can judge whether the use of a given video coding tool in a given local region of the video data yields a benefit or not. The encoder can set a flag indicating use of a tool at the level that encompasses the respective locality. To make this more tangible, consider a news program in which the presenters are seated in front of a static background, and introduce a news item concerning an event that comprises frenetic activity. When the newscast cuts away to footage of the event, the video characteristics change drastically – temporally and spatially. In the context of this matter, the requirement for loop filtering of smooth static background data, changes to a requirement for loop filtering of fast-moving detailed video objects. The video encoder will enable and disable types of filters, strengths of filters, and filter coefficients, on a local basis within the video frames according to the scene, and on a temporal basis as the sequence progresses. As another example, consider the adaption necessary during a football game between the seconds before a snap, and immediately afterward. It is thus typical for encoders to turn flags in the coded

bitstream on and off, dynamically at different levels of the data hierarchy. An added benefit is when a tool is disabled at a higher level, any syntax needed for the tool at a lower level can be omitted, thus reducing the volume of overhead information.

III. Relevant Legal Standards

85. I have been asked to provide my opinions as to whether claims 1–8 of the '556 Patent would have been obvious to a person of ordinary skill in the art at the time of the filing date of the '556 Patent.

86. 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 '556 Patent. In addition, I understand that the following legal principles apply.

87. It is my understanding that, in determining whether claims of the '556 Patent are unpatentable as 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.

88. 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.

89. 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 relevant 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.

90. 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.

91. 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 (but are not limited to): 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

92. It is my understanding that when interpreting the claims of the '556 Patent I must do so based on the perspective of one of ordinary skill in the art at the relevant priority date. My understanding is that the earliest claimed priority date of the '556 Patent is November 8, 2006.

93. It is my opinion that one of ordinary skill in the art in the field of video encoding and decoding, and more generally in the 2006 time period, would have had at least a bachelor's degree in electrical engineering, computer science, computer engineering, or a related field, and approximately three years of experience in design, development, and/or testing of relevant multimedia processing, such as

video, and/or image processing or coding. Advanced graduate studies with coursework and/or research focusing on topics of video encoding and decoding could substitute for working experience in the field.

94. I am a person of at least ordinary skill in the art and was so at the time of the purported invention of the '556 Patent.

V. Overview of the Subject Patent US 9,674,556, (“’556 Patent”)

95. The '556 Patent is titled “Methods and Apparatus for In-Loop De-Artifact Filtering.” It was filed on December 28, 2015 as U.S. Patent Application 14/981,345, which is a continuation of U.S. Patent Application 12/312,386 (“’386 Application, since issued as U.S. 9,227,243) originally filed as PCT/US2007/022795 on October 25, 2007. The '386 Application claims priority to U.S. Provisional Patent Application 60/964,917, filed on November 8, 2006, which for purposes of this declaration I assume to be the earliest priority date of the '556 Patent as I stated in paragraph 92, above.

A. Overview

1. Field of Art

96. The stated field of art is “video encoding and decoding,” specifically, “in-loop de-artifact filtering.” '556 Patent, 1:6–8.

2. Background - Problem statement, Admitted Prior Art

97. In the Background section, the '556 Patent describes two types of prior art filters and the known issues of applying each of those filters independently. First, the Background section describes prior art deblocking filters such as that implemented in the MPEG-4 AVC Standard. The Background admits that “[d]eblocking filtering in accordance with the MPEG-4 AVC Standard is well designed to reduce the blocky artifact but it does not try to correct other artifacts caused by quantization noise.” ’556 Patent, 1:58-61. Notably, later on the specification explains that “[t]he signal after [the MPEG-4] deblocking filter is a better match to the independent and identically distributed (i.i.d.) noise presumption.” *Id.*, 14:28-30.

98. The Background next describes a type of prior art filter that admittedly overcomes the limitations of MPEG-4 deblocking. *Id.*, 2:4-47. This filter is described as “a de-noising type nonlinear in-loop filter” that “adapts to non-stationary image statistics which exploits a sparse image model using an over complete set of linear transforms and hard-thresholding.” *Id.*, 2:6-10. The Background admits that this type of filter “automatically becomes high-pass, or low-pass, or band-pass, and so forth, depending on the region the filter is operating on” and “can address all types of quantization noise.” *Id.*, 2:13-14. The design of such a filter is stated to be: “This particular de-noising approach basically includes three steps: transform; transform coefficients threshold; and inverse transform. Then

several de-noised estimates provided by de-noising with an over complete set of transforms (typically produced by applying de-noising with shifted versions of the same transform) are combined using weighted averaging at every pixel.” *Id.*, 2:14-21. The Background section appears to have been written without a comprehensive review of the art for adaptive, sparse non-linear filters, because at 2:22-31, certain problems are stated and attributed to the prior art denoising tools being “designed for adaptive, independent, and identically distributed (i.i.d.) noise removal, while quantization noise has significantly different properties.” Afterwards, the Background states: “A possibility for a solution is spatio-frequential threshold adaptation, which may be able to correct the decision, but it is not trivial in its implementation.” *Id.*, 2:31-34. However, as discussed in the review of background art in this declaration, the prior art had already recognized this problem, and addressed it by such spatio-frequential adaptation. See for example the paper by Guleryuz.¹⁴

99. The Background section indicates that consequences of poorly designed sparsity based denoising tools include that “blocky artifact(s) may still be present despite the filtering procedure.” EX1001, 2:34-37. But after discussing the known MPEG-4 deblocking filter and known adaptive sparse de-noising filters, the

¹⁴ EX1010.

Background admits that “[i]t has been observed that the use of a single de-noising filter is not very efficient or effective in removing coding artifacts.” *Id.*, 2:48-50. It explains that “a general purpose de-noising filter is usually based on a distortion model which does not exactly match the actual scenario to which it is applied. This model does not consider the local structure of blocky artifact.” *Id.*, 2:50-54. It goes on to explain that a special purpose filter, such as the MPEG-4 deblocking filter “is not sufficient to correct the rest of the quantization noises. For example ... noise/artifacts at pixels away from the boundaries, within textures or to correct the distorted edges.” *Id.*, 2:54-62.

100. From the disclosure above, the '556 Patent established 1) the properties of the known MPEG-4 AVC Standard deblocking filter, 2) the properties of known adaptive sparse de-noising filters, and 3) the known problem that application of a single de-noising filter has shortcomings. I note the '556 discloses that there is no single filter that can resolve both the problem of noise on block boundaries and the problem of noise within blocks, notably on sharp edges inside blocks. But as amply disclosed in the overview of loop filtering for video data above, Section I.D, this situation had been recognized almost two decades earlier, and the use of two loop filters in succession had been publicly disclosed

3. Summary of the Alleged '556 Invention

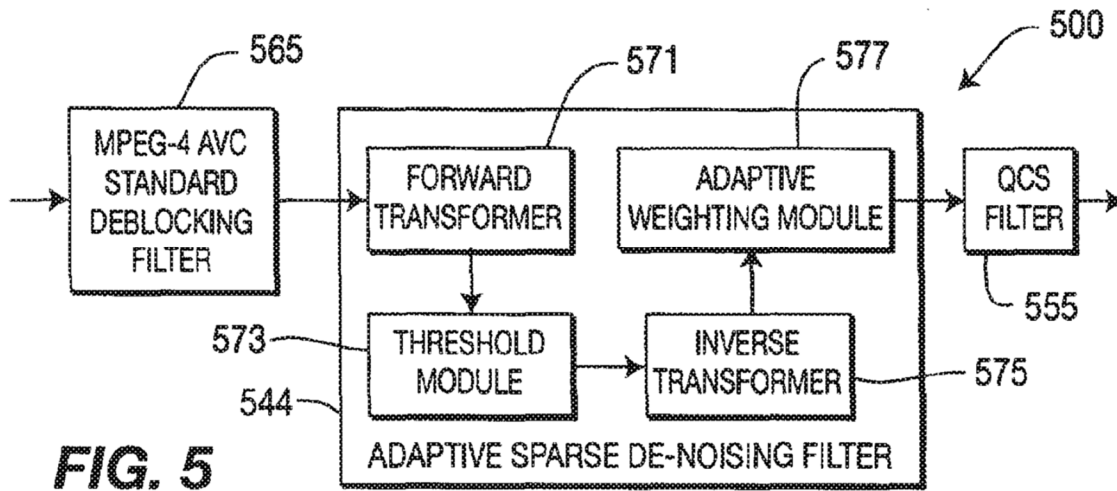
101. In brief, the alleged invention is summarized as a “method [] and apparatus for in-loop artifact filtering.” ’556 Patent, Abstract. Video compression artifacts result from quantization and are present in various forms. *Id.*, 1:12-16. The ’556 Patent describes the use of two or more in-loop filters to remove or reduce quantization artifacts. *Id.*, 10:17-29.

102. The asserted solution to the stated problem requires “utilizing a succession of filters on coded pictures to reduce the amount of coding artifacts over video frames.” *Id.*, 10:8-10. “The filters are placed in the coding loop (in-loop), so the filtered frames are used as referenced frames for motion compensation of subsequent coded frames.” *Id.*, 10:10-13. Finally, the specification states that in an embodiment, “the adaptive sparse de-noising filter ... can be turned on/off at the macroblock level, slice level, picture level, and/or sequence level.” *Id.*, 11:50-53.

4. Detailed Description

103. In the ’556 Patent, the use of two or more in-loop filters involves “the first filter (or filtering pass) to remove or reduce a first type (or group of types) of quantization artifact and the second filter (or filtering pass) to remove or reduce a

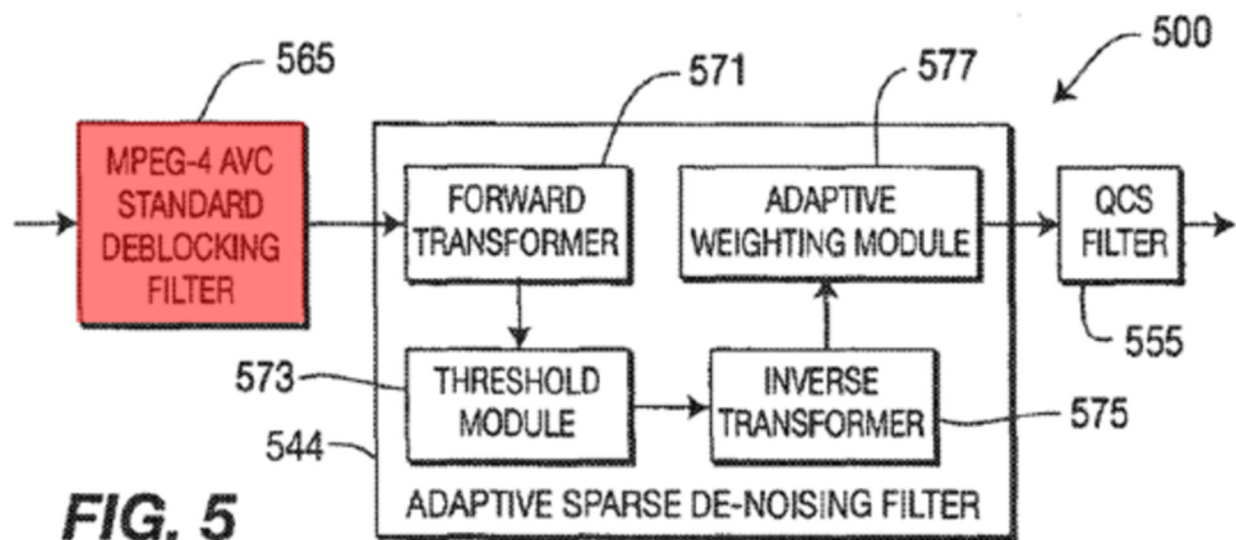
second type (or group of types) of quantization artifact.”¹⁵ *Id.*, 10:29-33. Figure 5 shows the use of the two or more in-loop filters:



104. In Figure 5, the deblocking filter, “which targets removing blocking artifacts,” is applied in the first “pass” and the adaptive sparse de-noising filter is applied in the second “pass” to “further reduce quantization noise over video frames composed of locally uniform regions (smooth, high frequency, texture, and so forth).” *Id.*, 10:8-13. However, I note that while the ’556 Patent provides descriptions of two filters—a deblocking filter and a sparse de-noising filter— it fails to provide details about how the proposed filters are different from the known (admitted) MPEG-4 deblocking filter and the known (admitted) adaptive sparse de-

¹⁵ The ’556 Patent explains that it uses the term “filtering passes” to refer to using a succession of filters. *Id.*, 10:42-45.

noising filter. In fact, the '556 Patent discloses that the deblocking filter *is* the MPEG-4 deblocking filter as I highlight (red) in Fig. 5 (565):



Id., Fig. 5 (annotated).

105. Figure 5 also illustrates that the admittedly known adaptive sparse denoising filter operates by applying thresholding in the transform domain. I note that nothing in the specification distinguishes the adaptive sparse denoising filter in Figure 5 from the known adaptive sparse denoising filter discussed in the background.

106. The '556 Patent further states that “[t]o further improve the result, we can enforce compression constraints including, but not limited to, quantization constraints, in order to help restore the image to the artifact-free form. ’556 Patent, 10:23-26. The ’556 Patent specifically teaches a QCS filter, applied after the

deblocking and “sparse de-noising” filters, for enforcing quantization constraints. E.g., *id.*, 11:6-16, Fig. 5, 15:24-58, Fig. 8. The ’556 Patent provides a specific example of a species of enforcing compression constraints, specifically quantization constraints, by using its QCS filter to “prevent over-smoothing” and “make[] the outcome of the cascade set of filters more robust to changes in the setting of filter,” *id.*, 11:11-16, “after artifact removal/reduction,” i.e., after application of its deblocking filter and “sparse de-noising filter.” *Id.*, 15:24-27. However, compression constraints comprise more than just quantization constraints, and include, e.g., pre-decided controls on filtering used by both the encoder and decoder to improve image quality. These include, e.g., the ’556 Patent’s QCS filter, as well as the flags controlling the ’556 Patent’s sparse filter. *Id.*, 12:1-50, Tables 1, 2. These are examples of “compression constraints” at least because their presence or absence limits or “constrains” the compression tools available to the encoder/decoder pair in order for the coded video to be usable.

107. The ’556 Patent also teaches that the adaptive sparse de-noising filter can be turned on/off at the macroblock, slice, picture, and/or sequence level. *Id.*, 11:50-53.

B. Claims

108. I list the Challenged Claims below. ’556 Patent, 18:24-19:8.

Claim 1:

A method comprising:

decoding an image region, wherein the decoding includes performing in-loop filtering to reduce at least a first and a second type of coding artifact using at least two filters in succession, the at least two filters including a deblocking filter for performing a first pass to reduce blocking artifacts and an adaptive sparse de-noising filter for performing a second pass to reduce noise, wherein the adaptive sparse de-noising filter is selectively enabled or disabled at a given level, the given level being at least a macroblock level, a slice level, a picture level or a sequence level.

Claim 2:

The method of claim 1, wherein at least one of the at least two filters is used to enforce at least one compression constraint.

Claim 3:

The method of claim 1, wherein at least one of the at least two filters is adaptively enabled or disabled depending upon at least data characteristics corresponding to the image region or data statistics corresponding to the image region.

Claim 4:

The method of claim 1, wherein a version of the image region that is displayed is different from at least one of an output signal corresponding to the image region after the at least two in-loop filters and an initial input signal corresponding to the image region before the at least two in-loop filters.

Claim 5:

An apparatus comprising at least a memory and one or more processors, the one or more processors being configured to:

decode an image region, the one or more processors including at least two filters for successively performing in-loop filtering to reduce at least a first and a second type of coding artifact, the at least two filters including a deblocking filter for performing a first pass to reduce blocking artifacts and an adaptive sparse de-noising filter for performing a second pass to reduce noise, wherein the adaptive sparse de-noising filter is selectively enabled or disabled at a given level, the given level being at least a macroblock level, a slice level, a picture level or a sequence level.

Claim 6:

The apparatus of claim 5, wherein at least one of the at least two filters is used to enforce at least one compression constraint.

Claim 7:

The apparatus of claim 5, wherein at least one of the at least two filters is adaptively enabled or disabled depending upon at least data characteristics corresponding to the image region or data statistics corresponding to the image region.

Claim 8:

The apparatus of claim 5, wherein a version of the image region that is displayed is different from at least one of an output signal corresponding to the image region after the at least two in-loop filters and an initial input signal corresponding to the image region before the at least two in-loop filters.

VI. Prosecution History of the '556 Patent

109. I have reviewed the prosecution history of U.S. Patent Application No. 12/981,345 (“the '345 Application”), from which the '556 Patent issued. From my review, I note that the '345 Application issued after minor, non-substantive amendments to its claims and the filing of a terminal disclaimer to mitigate undue patent term to the Applicant caused by the substantial overlap in claimed subject matter shared by the claims of the '345 Application and its parent '386 Application.

110. From my review, I noted the '386 Application was rejected by the Examiner during prosecution several times using prior art references to the same Yung-lyul Lee, et al. First, the Examiner rejected all claims as disclosed by Lee's

U.S. 6,665,346, specifically rejecting the application's claimed "sparse de-noising filter" as disclosed by Lee's deringing filter. '386 Application Prosecution History (EX-1005), 172-173.

111. Applicant amended the claims and argued the level in the video data hierarchy at which the filters were applied, appearing to admit that Lee taught two *types* of filters claimed, just not at the right *level*; Applicant amended the claims to enable and disable the filters on a slice and sequence level in addition to the previously-claimed block level. *Id.*, 189-90, 196. The Examiner then pivoted to another Lee reference (US Patent Application Publication 2003/0219073) to reject all the amended '386 claims again, and by re-mapping the deringing filter to the claimed "sparse de-noising filter," which at the time was recited in dependent claims. *Id.*, 281-82.

112. Applicant responded by changing the scope of the independent claims to require a second "pass" using the sparse de-noising filter. *Id.*, 287. Examiner then rejected the amended claims again, this time over the combination of Lee and applicant-admitted prior art ("AAPA"). *Id.*, 303-06. In this combination, the Examiner reasoned that it would have been obvious to combine Lee's teachings with the AAPA to "achieve a higher distortion reduction in terms of objective measures (e.g., mean squared error (MSE)) than other filters." *Id.*, 303-06. In my opinion, this makes excellent mathematical sense; reducing human-perceived distortion is a very

complex process that is difficult to quantitatively model, and MSE was a well-known model for accessing objective video quality (circa 1998) well preceding the time of the '556 Patent. A POSITA would have found it obvious to improve upon Lee's filter as a viable solution to reduce distortion. Sullivan, 78-79. I discuss the value of the AAPA teachings in more detail in Ground IB, below.

113. I see the Applicant responded to Examiner's position with an appeal, instead arguing the AAPA teaches away because it discloses a de-noising filter, but only to disparage it. '386 Application Prosecution History, 320-21. In my opinion, the Applicant's argument is without merit because, as I explain above, the Background section of the '556 Patent addresses issues with applying any one of the known prior art filters **independently** as the sole solution to the problem. Such an argument does not address the use of two or more filters in conjunction – which is the basis for the alleged invention – but in fact, this was well known in the art as amply demonstrated in this declaration. The prosecution history of the '386 Application does not contain any record of how or why the Examiner addressed this argument. The prosecution history of the '386 Application reflects only that Examiner and Applicant next had an interview and agreed upon adding “adaptive” to describe the “sparse de-noising filter” to overcome the rejections (*id.*, 333, 342-344) and achieve allowance, which Examiner issued without further elucidation (*id.*, 363-377).

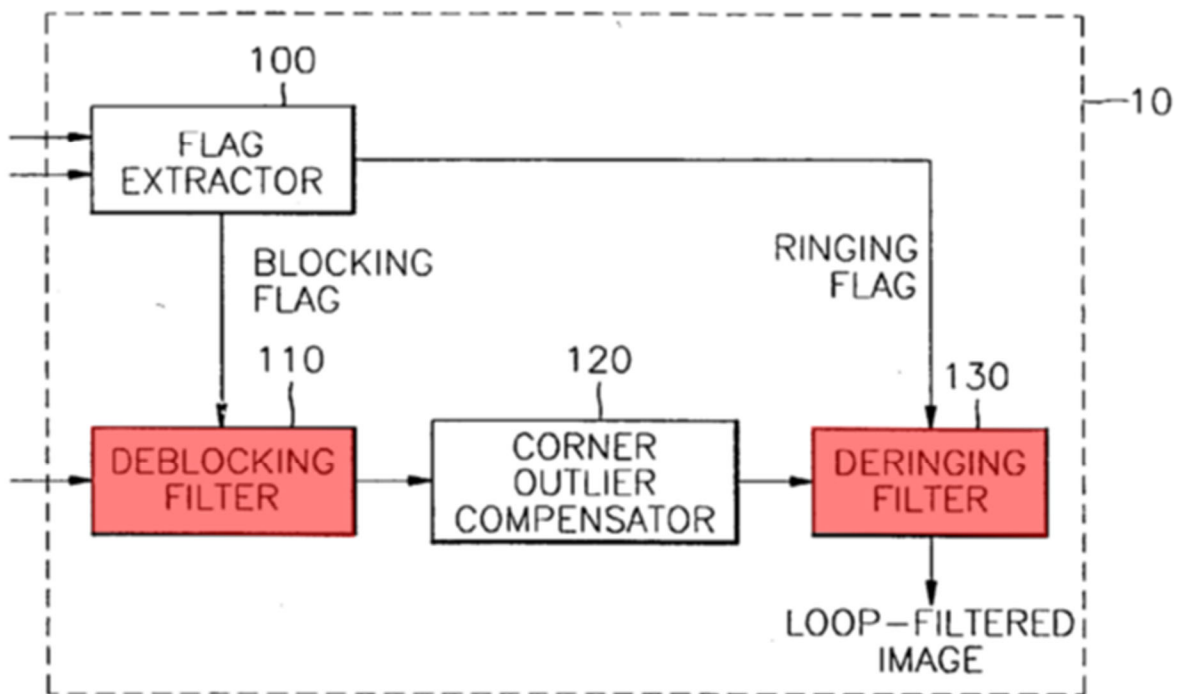
VII. Summary of Video Encoding and Decoding Prior Art

114. As I describe in more detail below, it is my opinion that the following pieces of prior art render obvious the alleged invention of the '556 Patent.

A. U.S. Patent Publication No. 2003/0219073 ("Lee")

115. Lee is directed to a loop-filtering method including a deblocking filter followed by a deringing filter. Lee, Abstract, [0044]. I have reproduced Lee's Figure 1 below, annotating each filter in red:

FIG. 1



116. In Figure 1, “an apparatus 10 for loop-filtering image data for reducing quantization effect occurring when the image data is block-base coded and decoded, which is used in connection with a general encoder and decoder, comprises a flag extractor 100, a deblocking filter 110, a corner outlier compensator 120 and a deringing filter 130.” Lee, ¶ 44. Lee discloses the use of flags on a block level, with deringing filtering enabled or disabled based on a deringing flag. *See id.*, ¶¶ 10-12, 24, 118.

B. Applicant Admitted Prior Art (“AAPA”): Adaptive sparse de-noising filter

117. The ’556 Patent admits that adaptive sparse de-noising filters are prior art. Specifically, the ’556 Patent admits that in order to overcome the limitations of the MPEG-4 AVC Standard deblocking filter, an approach has been “recently” proposed involving a de-noising type nonlinear in-loop filter. In this proposed approach, a nonlinear de-noising filter adapts to non-stationary image statistics which exploits a sparse image model using an over complete set of linear transforms and hard-thresholding. The nonlinear de-noising filter automatically becomes high-pass, or low-pass, or band-pass, and so forth, depending on the region the filter is operating on. ’556 Patent, 2:4-13.

118. It is clear that the ’556 Patent discloses a prior art “adaptive sparse de-noising filter” because it: (1) “adapts to non-stationary image statistics,” i.e., based

on the region being analyzed; (2) uses a mathematically sparse filter design; and (3) is intended to reduce or eliminate noise. The '556 Patent further goes on to explain that “the [adaptive sparse de-noising filter] was designed for additive, independent and identically distributed (i.i.d.) noise removal.” '556 Patent, 2:25-27.

119. The '556 Patent claims that the referenced sparse de-noising filter, when applied on its own, has known drawbacks such as being “not trivial in its implementation” and “present[ing] important visual artifacts that need to be addressed,” but it nonetheless admits that known sparse de-noising filters “present a higher distortion reduction in terms of objective measures (e.g., mean square error (MSE)) than other techniques.” *Id.*, 2:31-47.

C. Applicant Admitted Prior Art (“AAPA”): Standard deblocking filter

120. The '556 Patent admits that “[a]daptive deblocking filters have been studied in the past and some well-known deblocking filtering methods have been proposed and adopted in various standards.” '556 Patent, 1:23-25. The specification then focuses on prior art in-loop deblocking filtering in accordance with the MPEG-4 AVC Standard. *E.g.*, '556 Patent, 1:33-2:21, 2:63-4:54, Figs. 1-3. The specification explains that “[d]eblocking filtering in accordance with the MPEG-4 AVC Standard is well designed to reduce the blocky artifact, but it does not try to correct other artifacts caused by quantization noise.” *Id.*, 1:58-61. The specification further notes

that “quantization noise in video compression, and particularly with respect to the MPEG-4 AVC Standard, is usually not independent and identically distributed (i.i.d.).” *Id.*, 14:22-24. However, “[t]he signal after [MPEG-4 AVC Standard] deblocking filter is a better match to the independent and identically distributed (i.i.d.) noise presumption.” *Id.*, 14:28-30.

D. U.S. Patent Publication No. 2006/0110062 (“Chiang”)

121. Chiang teaches “an edge adaptive filtering system [and method] for reducing artifacts” in a decoded image. Chiang, ¶ 5. Chiang’s image decoder implements an edge adaptive filter in conjunction with a de-blocking filter. Chiang, ¶ 34. Chiang’s edge adaptive filter “processes image information to remove ringing artifacts, Chiang, ¶ 41. Chiang’s deblocking filter reduces blocking artifacts. Chiang, ¶ 29. Chiang describes that each of the de-blocking filter and the edge adaptive filter can be implemented in-loop. Chiang, ¶¶ 34, 40.

122. In Chiang, the adaptive edge filter adaptively filters image information in a block by using the filter weights identified from a visual weight map generator. Chiang, ¶ 56. The visual weight map generator identifies filter weights based on input from a texture analyzer, an edge analyzer, a quality analyzer, and a monotone detector, each of which analyze portions of the image being decoded. Chiang, ¶ 55. In some circumstances, Chiang recognizes that it is appropriate to set filter weights to 0, meaning that no adaptive edge filtering occurs. Chiang, ¶ 81.

E. U.S. Patent Publication No. 2003/0152146 (“Lin”)

123. Lin provides techniques and tools for processing reference frames in a motion estimation/compensation loop or motion compensation loop. Lin, Abstract. Lin teaches its decoder using one or more in-loop filters, and that in-loop filtering can be enabled or disabled on “a per-frame, per-sequence, or other basis.” Lin, ¶¶ 124, 126, 131.

F. SMPTE Standard: VC-1 Compressed Video Bitstream Format and Decoding Process, SMPTE 421M-2006 (“VC1 Standard”)

124. The VC1 Standard is an SMPTE (the Society of Motion Picture and Television Engineers) standard based on WMV9 (Windows Media Video 9) developed by Microsoft. The VC1 Standard contains coding tools for interlaced video sequences as well as progressive encoding.

125. The VC1 Standard describes the decoding process for decoding progressive pictures and interlace pictures. VC1, 120, 253. Two filtering operations, i.e., overlap smoothing and in-loop deblocking, are included in a decoding process. E.g., VC1, 4, 5, Figs. 3, 4. Relevant here, the in-loop deblocking filtering is performed to smooth out the discontinuities at block boundaries. VC1, 178. In addition, two filtering operations, i.e., deblocking filtering and deringing filtering, are included in the post-processing. VC1, 452.

126. Syntax elements are included in a bitstream to be decoded that selectively enable or disable the filtering operations. VC1, 463. These syntax

elements include a LOOPFILTER flag to control the in-loop deblocking filtering, VC1, 462, an OVERLAP flag to control overlap smoothing, VC1, 463, and a POSTPROC flag to control the two post-processing filtering operations, VC1, 452.

G. Technical Article: “A Nonlinear Loop Filter for Quantization Noise Removal in Hybrid Video Compression” (“Guleryuz”)

127. Guleryuz is a technical article that discusses the use of loop filters for video compression. Guleryuz, 1. Guleryuz discloses a “high-performance, nonlinear, loop filter that reduces quantization noise over video frames composed of locally uniform regions (smooth, high frequency, texture, etc.) separated by singularities.” *Id.*, Abstract. Guleryuz explains that its filter “is not limited to block based transform coders and provides a robust solution for intra as well as differentially encoded frames.” *Id.* Guleryuz’s filter reduces quantization noise. *Id.*, 1-2. Such noise includes both blocking artifacts and artifacts around sharp edges. *Id.*, 3-4.

VIII. Claim Construction

128. Generally, the Challenged Claims need no construction to demonstrate they are unpatentable. I understand that in related § 337 investigation before the International Trade Commission, the parties and ITC Staff present four competing constructions for the term “sparse de-noising filter” as recited in claims 1 and 5. The table below summarizes the proposed constructions:

Complainant (Patent Owner)’s Construction	Respondents’ Construction	Staff’s Construction
---	---------------------------	----------------------

Plain and ordinary meaning (i.e., a filter that reduces noise based on a sparse representation of the signal)	Sparse de-noising filter for performing a second pass to reduce noise, which is not a deringing filter Alternatively, a filter which exploits a sparse image model using an over complete set of linear transforms and hard thresh-holding, which is not a deringing filter	A filter that reduces noise based on a sparse representation of the signal
---	--	--

EXs1025, 1032-1033

129. I first note that although Patent Owner and the Staff appear to propose the same construction in the related ITC proceeding, I understand that there is a dispute as to its application. Specifically, as seen from the excerpt from the Staff's *Markman* brief below, the dispute is whether PO's interpretation, which encompasses sparse signals as well as sparse representations of signals, and is apparently driven by its infringement argument, is correct:

¹⁰ Although the Staff agrees with Complainants' construction, there is a dispute among the parties regarding whether the signal itself can be sparse (i.e. whether "a sparse representation of the signal" encompasses a sparse signal). The Staff is of the view that "a sparse representation of the signal" in its proposed construction *does not* encompass a sparse signal (i.e., the representation of the signal is sparse, not the signal itself). The Staff understands that Complainants disagree and that the dispute is based on infringement arguments related to the functionality of the accused products. In the Staff's view, the evidence and, in particular, the extrinsic evidence supports its understanding.

EX1025, Staff *Markman* Br., n.10. Thus, even assuming that a "sparse de-noising filter" is construed to mean "a filter that reduces noise based on a sparse representation of the signal," I understand there is a dispute whether the construction

encompasses that the signal itself is sparse, or whether just the representation of the signal is sparse.

130. I also note the inconsistent positions taken by Patent Owner with respect to “sparse de-noising filter.” Specifically, I understand that during prosecution before the USPTO (as I detail above), the applicant argued the “sparse de-noising filter” does not encompass a “deringing” filter, but in the related ITC proceedings, PO now argues the term “sparse de-noising filter” *is* met by a filter that “perform[s] deringing.” EX1024, 7, 10; EX1031, 10. This introduces a question of scope – whether the term “sparse de-noising filter” excludes a “deringing” filter (which I understand is Petitioner’s position in the Petition here),¹⁶ or is broad enough to encompass any “filter that reduces noise based on a sparse representation of the signal” (as Patent Owner alleged before the ITC). I also understand that beyond this dispute, and as reflected by Respondent’s alternative construction in the ITC proceeding, there is further dispute regarding whether it is necessary to construe the term further, and if necessary, the proper construction for the term. I address these specific questions in the paragraphs following, but also have the opinion that

¹⁶ I understand that Petitioner is not a party to the co-pending ITC proceeding, but that Petitioner has identified the ITC Respondent as a real-party-in-interest to this IPR proceeding.

regardless of which is the correct construction, the prior art renders obvious the Challenged Claims.

131. To be sure, one of the two filters recited in the independent claims is a “sparse de-noising filter.” In my opinion, a “sparse de-noising filter” is a filter that conducts filtering (i.e., “de-noising”) by utilizing a “sparse” mathematical model. This is consistent with the applicant’s arguments during prosecution of the ’386 Application and Respondents’ position taken in the related ITC proceeding because the deringing filter disclosed by Lee, while reducing ringing artifacts across image edges, does not utilize a sparse mathematical model to do so. It is also consistent with ITC Respondents’ alternative construction and the disclosure of the ’556 Patent, in which the “sparse de-noising filter” conducts filtering by determining an “over complete set of linear transforms and hard-thresholding” set for a given level of image information. ’556 Patent, 2:9-10, 13:24-34. This operation results in a mathematically sparse image model (see the discussion of the Li/Orchard prior art in the Video Encoding and Decoding Technologies section above, ¶¶ 73-74). This is also consistent with the plain English understanding of the phrase “sparse de-noising filter.” The word “sparse” is an adjective limiting the scope of the noun filter. So it is the filter that is “sparse,” as opposed to the data on which the filter operates being sparse. Notably, the ITC Staff’s proposed construction is more consistent with, though slightly broader than, Respondents’ alternative construction than it is with

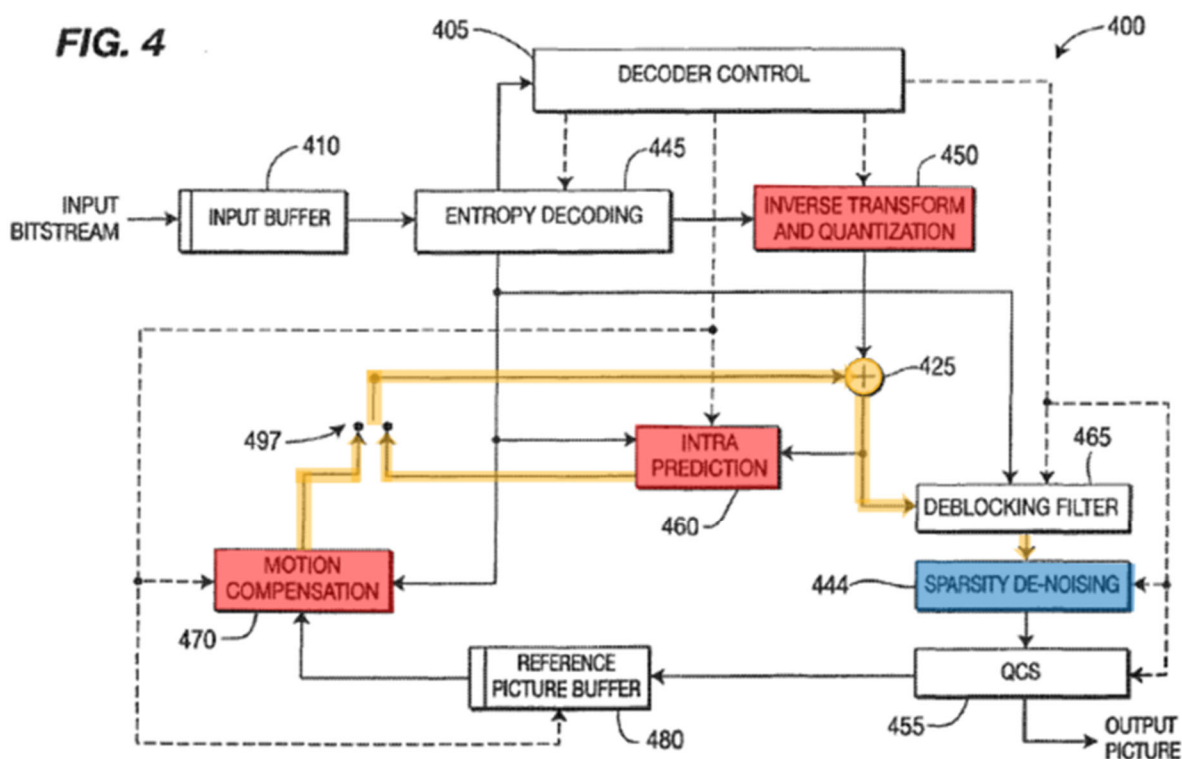
PO's construction. Both constructions contemplate a filter that operates on data that is not sparse, the filter transforming and filtering the signal using a sparse mathematical model, as I explain above. The only substantive difference between these constructions, aside from the exclusion of deringing filters by Respondents' construction, is that the Staff's construction is broader in the sense that it does not proscribe how sparsity is achieved by the recited filter, while Respondents' alternative construction uses the '556 Patent's own description of the filter to do so.

132. From my review, there is no basis for Patent Owner's ITC construction (as Patent Owner apparently interprets that construction)—that the filter “reduces noise based on a sparse representation of the signal” and a “sparse representation of the signal” encompasses a sparse signal. EX1025, n.10; EX1024, 7, 9. Indeed, it appears that PO reads out the word “representation” from its own construction. In my opinion, “a representation of the signal” is not the signal itself. I see no support for PO's construction and interpretation in the intrinsic record.

133. First, the term “sparse representation” appears only once in the specification to describe a prior-art filter with inadequate threshold values (which are used to trigger when the filter is applied), and results in a signal with “sparse representation at the filtering stage.” '556 Patent, 2:37-41. This describes the noise-based attributes of a *signal*, which is wholly separate and unrelated to the claimed “sparse de-noising *filter*” used to reduce the noise within an image.

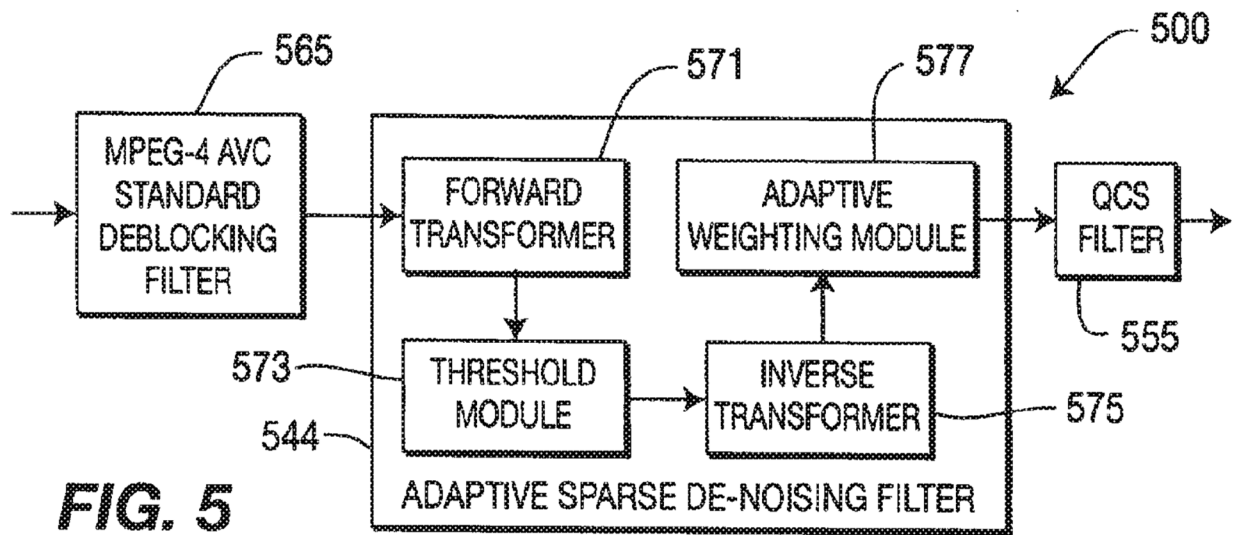
134. Second, PO's ITC construction is undermined by Figure 4 and its related discussion. Figure 4 (which I annotated and reproduced below) depicts a decoder applying the alleged invention of the '556 Patent and is reproduced below.

'556 Patent, 9:8-12.



135. This particular arrangement of components—having the “sparsity de-noising” filter (blue) receive a signal that has already been de-quantized (by the “inverse transform and quantization,” 450) (red) and combined (425) (orange) with other compensators (470) (red) and predictors (460) (red)—would have conveyed to a POSITA that the series of filters (e.g. the deblocking, de-noising, and QCS filters) would not be receiving “a sparse representation of a signal” at all.

Specifically, signal transformation and quantization are performed to create a sparse representation of a signal (e.g., for transmission), and inverse transformation and quantization reverses this processes, as the name suggests, resulting in a dense (i.e., not sparse) signal. This understanding is supported by the '556 Patent's Figure 5, reproduced below:



136. Although it is my opinion that the ITC Respondent's proposed construction and alternative construction are correct, for the purposes of this proceeding I have been instructed to analyze Petitioner's grounds under each of the four proposed constructions in the ITC proceeding. As detailed in the grounds below, it is my opinion that the challenged claims are unpatentable as obvious regardless of the construction adopted by the Board.

IX. Ground IA – Claims 1–8 are Obvious under 35 U.S.C. § 103 by Lee

137. In my opinion, based upon PO’s current position with the ITC that a “sparse denoising filter” is a “filter that reduces noise based on a sparse representation of the signal,” which encompasses a sparse signal, claims 1–8 of the ’556 Patent rendered obvious by Lee. In the discussion of claim 1 that follows, I explain the correspondence between the disclosure of Lee to each limitation of claim 1 under PO’s current ITC claim construction.

A. Claim 1

138. The preamble of claim 1[pre] recites: “**A method comprising:**”

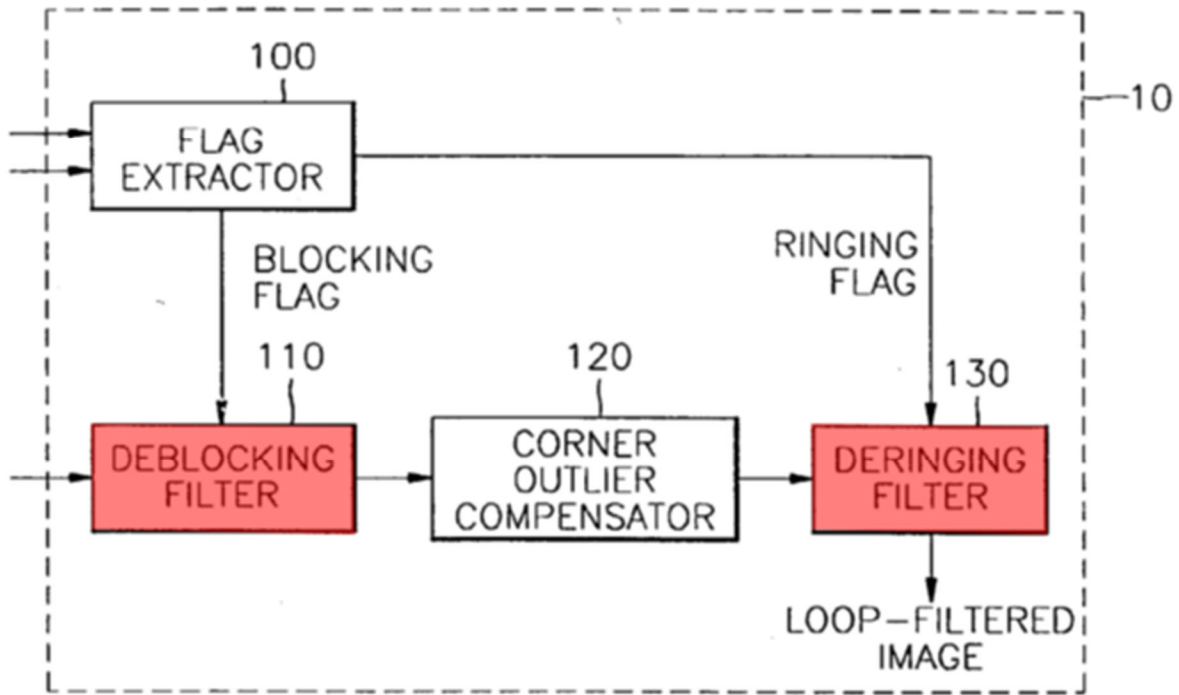
139. The prior art discloses or suggests this, to the extent the preamble is limiting, as follows. Lee discloses “[a] loop-filtering method for reducing quantization effect generated when an image data is encoded and decoded.” Lee, Abstract.

140. The first limitation of claim 1[a1] recites: “**decoding an image region, wherein the decoding includes performing in-loop filtering to reduce at least a first and a second type of coding artifact using at least two filters in succession,**”

141. The prior art discloses or suggests this step as follows. Lee discloses “loop-filtering image data for reducing quantization effect occurring when [] image data is block-base coded and decoded, which is used in connection with a general encoder and decoder.” Lee, ¶ 44. This is accomplished using the apparatus depicted

in Lee's Figure 1, which I have annotated below to highlight the deblocking and deringing filters:

FIG. 1



142. Lee explains that “[t]he basic idea of the loop-filtering method ... is to adaptively filter using spatial frequency and temporal information to reduce the quantization effect.” Lee, ¶ 50. Therefore, Lee discloses “decoding an image region, wherein the decoding includes performing in-loop filtering.”

143. Lee's loop-filtering reduces at least two types of coding artifacts – blocking and ringing artifacts – using at least two filters in succession – a deblocking

filter and a deringing filter. In particular, Lee teaches using its deblocking filter 110 to reduce a first type of coding artifacts: blocking artifacts. Lee, ¶¶ 70-72. The deblocking filter 110 checks for a blocking flag extracted by block extractor 100 to determine whether to perform deblocking. Lee, ¶ 47.

144. Further, corner “outlier compensator 120 detects the corner outlier of the data passed through the deblocking filter 110, and compensates for the detected corner outliers.” Lee, ¶ 48. Lee explains “[c]orner outlier compensation is only performed for INTRA frame.” *Id.*, ¶ 100. Therefore, the corner outlier compensator is not applied when decoding an inter frame.

145. Additionally, Lee’s deringing filter 130 reduces a second type of coding artifact – ringing artifacts. Lee, ¶¶ 70-71, 104. “The deringing filter 130 checks the ringing flag extracted by the flag extractor 100 ... to deringing-filter the corner outlier compensated data.” Lee, ¶ 49. In the case of decoding an inter frame, the ringing flag check and deringing filtering (if indicated by the flag) is done immediately after the deblocking filtering because the intervening corner outlier compensator is not applied. Lee, ¶ 104. Therefore, Lee teaches using its deblocking and deringing filters in succession because deringing is performed after deblocking and corner compensation, or immediately after deblocking when decoding inter frames.

146. The second limitation of claim 1[a2] recites: **“the at least two filters including a deblocking filter for performing a first pass to reduce blocking artifacts and an adaptive sparse de-noising filter for performing a second pass to reduce noise,”**

147. The prior art discloses or suggests this structure as follows under PO’s current position with the ITC that a “sparse denoising filter” is a “filter that reduces noise based on a sparse representation of the signal,” which encompasses a “sparse signal.” As I discuss above, it is my opinion that PO’s position as to the construction of this term is incorrect. Nevertheless, to the extent the Board adopts PO’s construction, Lee discloses this limitation. As I explained in paragraph 139, Lee teaches using a deblocking filter 110 to reduce blocking artifacts. Lee, ¶¶ 70-72. Lee also teaches that the deblocking filter is applied before its corner outlier compensator 120 and deringing filter 130. Lee, ¶¶ 47-49, Fig. 1. Therefore, deblocking filter 110 performs a first “pass” of filtering as the term is used in the ’556 Patent.

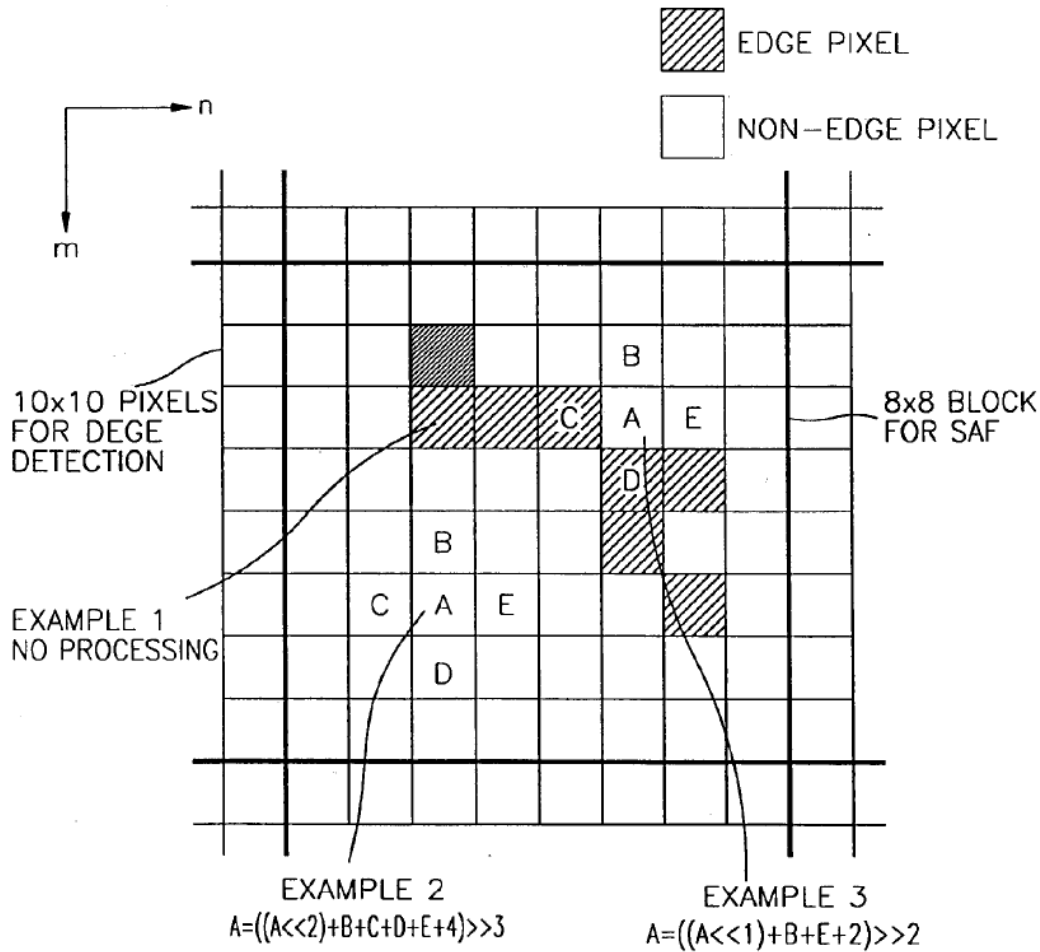
148. As I also explained in paragraphs 141–145, Lee further teaches using a deringing filter 130 after the first pass of deblocking filter 110. Lee, ¶¶ 70-71, 104. Under PO’s apparent interpretation of “sparse de-noising filter,” Lee’s deringing filter is a “sparse de-noising filter.” In particular, Lee’s deringing filter reduces “ringing noise near image edges.” Lee, ¶¶ 5, 12, 31, 45. This is consistent with PO’s application of the term to the CDEF filter in the AV1 standard, which PO claims to

be a “sparse de-noising filter” because it “prevent[s] excessive blurring when the filter is applied across an edge” and therefore is “adapts to the sparse nature of underlying video signals.” EX1024, 9; EX1031, 10. Lee’s deringing filter is an “adaptive sparse de-noising filter” under PO’s interpretation because it 1) creates an “edge map” of an image (“a sparse representation of a signal”) and 2) performs filtering “based on” said edge map. To create an edge map, Lee’s deringing filter detects horizontal and vertical edges by comparing the differences between neighboring pixels to threshold values, which are based on the H.263 “quantization factor” or QP. Lee, ¶ 24. If the difference between two neighboring pixels is large (i.e., greater than the threshold value), Lee determines that pixel to be an image edge pixel and sets the edge map value at that location to 1. *Id.* If the difference is small, Lee determines the pixel is not an edge pixel, and sets the edge map value to 0. *Id.*, ¶¶24, 111-113, Fig. 7B. Lee’s paragraph 24 summarizes the specific calculations employed:

In the step (a) for the horizontal edge detection, assuming that a pixel[m,n] is the current pixel, a pixel[m][n+1] is located to the right of the pixel[m,n], a pixel[m][n-1] is located to the left of the pixel[m,n], the difference between the pixel[m][n] and the pixel[m][n+1] is A1, the difference between the pixel[m][n] and the pixel[m][n-1] is A2, and QP is the H.263 quantization factor, when the conditional expression ((A1>QP) and (A2>QP)) or (A1>2QP) or (A2>2QP) is satisfied, the current pixel is detected as an edge pixel, and the edge map, Edge[m][n], becomes 1. Also, in the step (a) for the vertical edge detection, assuming that the current pixel is pixel[m][n], a pixel[m+1][n] is located above the pixel[m][n], the lower pixel is pixel[m-1][n], the difference between the pixel[m][n] and the pixel[m+1][n] is A'1, the difference between the pixel[m][n] and the pixel[m-1][n] is A'2, and QP is the H.263 quantization factor, when the conditional expression ((A'1>QP) and (A'2>QP)) or (A'1>2QP) or (A'2>2QP) is satisfied, the current pixel is detected as an edge pixel, and the edge map, Edge[m][n], becomes 1.

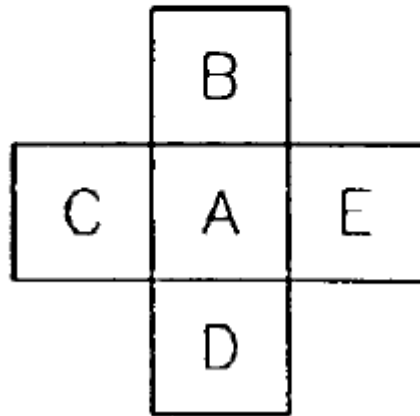
The resulting edge map is made up of entirely ones and zeros that represent the location of the edges in the original image. As Lee's Figure 7B demonstrates, this representation is "sparse" because relatively few pixels (just the edge pixels) have a value of one and the majority of pixels have a zero-value.

FIG. 7B



149. Lee then filters adaptively “based on” the edge map. Lee, ¶¶24, 113. Lee’s deringing filter analyzes 4-connected filter windows to determine if the filter windows contain an edge pixel. Lee, ¶113. These windows are depicted in Lee’s Figure 7A:

FIG. 7A



If the center pixel of the filter window is an edge pixel, no filtering is performed. *Id.* If a non-center pixel is an edge pixel, Lee performs weighted filtering to exclude edge pixels. *Id.* And if no edge pixel is included, low-pass filtering is performed. *Id.* Lee's deringing filter therefore performs filtering "based on" the edge map ("sparse representation of a signal") it creates. This filtering is also therefore adaptive because the filtering weights adapt based on the edge map. *Id.*, ¶¶ 24, 50 ("The basic idea of the loop-filtering method on the image data according to the invention is to adaptively filter using spatial frequency and temporal information to reduce the quantization effect."), 111-113.

150. And as I explained in paragraphs 141–145, in decoding inter frames, Lee's deringing filter is applied after the deblocking. Lee, ¶ 100. Therefore, the

application of Lee's adaptive deringing filter is "performing a second pass to reduce noise."

151. The third limitation of claim 1[a3] recites: **"wherein the adaptive sparse de-noising filter is selectively enabled or disabled at a given level, the given level being at least a macroblock level, a slice level, a picture level or a sequence level."**

152. The prior art discloses or suggests this step as follows. Lee teaches "extracting a flag indicating whether the image data requires loop filtering." Lee, ¶¶ 10, 57. Lee further teaches that this flag "comprises ... a ringing flag indicating whether the reduction of ringing noise near image edges is required." *Id.*, ¶ 12. In particular, "the deringing filtering for reducing the ringing noise in the filtering step is performed if the ringing flag indicates a need for the deringing filtering, and is not performed if the ringing flag indicates no need for the deringing filter." *Id.*, ¶ 24. Lastly, Lee teaches that its adaptive deringing filter flag is applied at the macroblock level for filtering inter frames. *Id.*, ¶¶ 63 ("Propagation of Flag Information for Inter-Macroblock (Inter MB)"), 55 ("The blocking and ringing flags are extracted from the DCT domain of each 8x8 block in the intra-macroblock (MB). ***Also, the flags of the inter-MB*** are calculated from a residual signal and flags of a reference video object plane (VOP)") (my emphasis added). Thus, it is my opinion that Lee's ringing

flag “selectively enable[s] or disable[s] at a given level, the given level being at least a macroblock level.”

B. Claim 2

153. Claim 2 recites: **“The method of claim 1, wherein at least one of the at least two filters is used to enforce at least one compression constraint.”**

154. The prior art discloses this step under PO’s current position with the ITC regarding “sparse de-noising filter” as follows. Lee discloses at least two types of compression constraints enforced by Lee’s deblocking filter.

155. First, Lee discloses “extracting a flag indicating whether the image data requires loop-filtering using the distribution of inverse quantized coefficients (IQCs) of an inverse quantized image data and a motion vector indicating the difference between the previous frame and the current frame.” Lee, ¶ 10. Lee further discloses that this flag comprises, *inter alia*, “a blocking flag indicating whether reduction of blocking artifacts near block boundaries is required.” *Id.*, ¶ 12, *see also, id.*, ¶¶ 31, 45. That indication of the deblocking flag from the inverse quantized coefficients is used by Lee’s deblocking filter to enforce a compression constraint because both the encoder and decoder are constrained ahead of time to use syntax and tools available, in this case, the deblocking filter as taught by Lee, to ensure that which is encoded can later be decoded.

156. Second, Lee discloses several possible rules for performing its deblocking filtering based on extracted blocking flags and ringing flags (to control the deringing filter). *E.g., id.*, ¶¶ 20-22, 76, 88. For example, Lee discloses a circumstance where, a first filtering along a block boundary is performed and then, considering two adjacent blocks I and J, if the ringing flags for both indicate no need for deringing filtering and the horizontal blocking flags indicate a need for reducing blocking artifact, “comparing the difference between two adjacent pixels around the block boundary with the H.263 quantization factor QP” and performing a second filtering if certain flags remain and “if the difference between two pixels is less than the QP.” *Id.*, ¶ 20, *see also, id.*, ¶¶ 21-22. In other words, Lee discloses its deblocking filter filtering once, reassessing by comparing pixel value differences along adjacent block boundaries to a predetermined value, and filtering again as needed. The blocking filter thus enforces a quantization constraint, which is a subset of the broader category of compression constraints, by comparing the quantization along block boundaries to the H.263 quantization factor and filtering until the image is sufficiently improved.

157. Accordingly, Lee discloses at least two different manners in which its deblocking filter enforces at least two different rules that are encompassed within the term “compression constraint.” First, the use of the blocking flag, whose presence or absence indicates to the decoder whether to apply the deblocking filter,

is enforced by the deblocking filter. And second, the deblocking filter performs multiple rounds of deblocking filtering, as needed, by comparing the pixel values along a block boundary against the H.263 quantization factor to determine whether blocking artifacts have been sufficiently filtered out.

C. Claim 3

158. Claim 3 recites: **“The method of claim 1, wherein at least one of the at least two filters is adaptively enabled or disabled depending upon at least data characteristics corresponding to the image region or data statistics corresponding to the image region.”**

159. The prior art discloses this step under PO’s current position with the ITC regarding “sparse de-noising filter” as follows. Lee teaches “the deringing filtering for reducing the ringing noise ... is performed if the ringing flag indicates a need for the deringing filtering, and is not performed if the ringing flag indicates no need for the deringing filtering.” Lee, ¶ 24.

160. Lee further teaches that the value of its ringing flag depends on various image characteristics. For example, Lee teaches that deringing filtering is performed when “the block I is an intra block, when the block I is not an intra block and has an 8×8 prediction mode, and when the block I is not an intra block and has no 8×8 prediction mode and the RF 1 of the block 1 indicates no need for the deringing

filtering,” and “is not performed when the RF1 of the block I indicates no need for the deringing filtering.” *Id.*, ¶ 27.

161. With even more particularity, Lee teaches the circumstances under which its deringing filter is applied when decoding an inter macroblock. *Id.*, ¶¶ 63, 66-69. Lee teaches that “the RF [ringing flag] of the reference block A c 0 is set to 1 if the IQC of the residual signals in the 8×8 block of the inter-frame is non-zero.” *Id.*, ¶ 66. Lee also teaches setting RF to 1 if four motion vectors are transferred from one macroblock. *Id.* In another embodiment, which I note solely for the purposes of demonstrating that Lee broadly teaches application of its ring flag, Lee teaches setting the ringing flag based on the IQC (inverse quantized coefficients) of macroblock of interest. *Id.*, ¶ 69. In any case, the ringing flag RF being set to 1 adaptively enables the deringing filter based on the data characteristics or data statistics of the macroblock, i.e., “image region,” of interest.

D. Claim 4

162. Claim 4 recites: “**The method of claim 1, wherein a version of the image region that is displayed is different from at least one of an output signal corresponding to the image region after the at least two in-loop filters and an initial input signal corresponding to the image region before the at least two in-loop filters.**”

163. The prior art discloses or suggests this step under PO's current position with the ITC regarding "sparse de-noising filter" as follows. Lee discloses that "the version of the image region that is displayed is different from ... an initial input signal corresponding to the image region before the at least two in-loop filters." Lee explains that its deblocking and deringing filters "reduce[] the quantization effect" "such as blocking artifacts ... and ringing noise." Lee, Abstract, ¶ 118. It follows that the reduction of the blocking and ringing artifacts by Lee's deblocking and deringing filters results in a "version of the image region that is displayed [being] different from ... an initial input signal corresponding to the image region before the at least two in-loop filters." Lee, ¶¶ 6-7 (discussing effect of deblocking and deringing filters on images after decoding).

E. Claim 5

164. Claim 5[pre] recites: "**An apparatus comprising at least a memory and one or more processors, the one or more processors being configured to:**"

165. To the extent it is limiting, the prior art discloses or suggest this structure as follows. Lee discloses that its teachings "may be embodied in a general purpose digital computer that is running a program from a computer usable medium, including but not limited to storage media such as magnetic storage media (e.g., ROMs, floppy disks, hard disks, etc.), optically readable media (e.g., CD-ROMs, DVDs, etc.) and carrier waves (e.g., transmissions over the Internet)." Lee, 116. It

was well known at the relevant time that a general purpose digital computer comprises one or more processors and that the examples of “computer useable medium” recited by Lee comprise “a memory.”

166. The first limitation of the claim, **5[a1]**, recites: **“decode an image region, the one or more processors including at least two filters for successively performing in-loop filtering to reduce at least a first and a second type of coding artifact,”**

167. The prior art discloses or suggests this limitation for the same reasons as I explained in paragraphs 141–145.

168. The second limitation of the claim, **5[a2]**, recites: **“the at least two filters including a deblocking filter for performing a first pass to reduce blocking artifacts and an adaptive sparse de-noising filter for performing a second pass to reduce noise,”**

169. The prior art discloses or suggests this structure under PO’s current position with the ITC for the same reasons as I explained in paragraphs 147–150.

170. The third limitation of the claim, **5[a3]**, recites: **“wherein the adaptive sparse de-noising filter is selectively enabled or disabled at a given level, the given level being at least a macroblock level, a slice level, a picture level or a sequence level.”**

171. The prior art discloses or suggests this structure under PO's current position with the ITC for the same reasons as I explained in paragraph 152.

F. Claim 6

172. Claim 6 recites: **"The apparatus of claim 5, wherein at least one of the at least two filters is used to enforce at least one compression constraint."**

173. Lee discloses or suggests this structure under PO's current position with the ITC regarding "sparse de-noising filter" for the same reasons as I explained in paragraphs 154-157.

G. Claim 7

174. Claim 7 recites: **"The apparatus of claim 5, wherein at least one of the at least two filters is adaptively enabled or disabled depending upon at least data characteristics corresponding to the image region or data statistics corresponding to the image region."**

175. The prior art discloses or suggests this structure under PO's current position with the ITC regarding "sparse de-noising filter" for the same reasons as I explained in paragraphs 159–161.

H. Claim 8

176. Claim 8 recites: **"The apparatus of claim 5, wherein a version of the image region that is displayed is different from at least one of an output signal corresponding to the image region after the at least two in-loop filters and an**

initial input signal corresponding to the image region before the at least two in-loop filters.”

177. The prior art discloses or suggests this structure under PO’s current position with the ITC regarding “sparse de-noising filter” for the same reasons as I explained in paragraph 163.

X. Ground IB – Claims 1–8 Are Obvious under 35 U.S.C. § 103 Based on Lee in View of AAPA

178. In my opinion a person of ordinary skill in the art in 2006 would have found claims 1–8 of the ’556 Patent to have been obvious based on Lee in view of AAPA under any of the proposed constructions for “sparse de-noising filter.” The AAPA “exploits a sparse image model using an over complete set of linear transforms and hard-thresholding,” ’556 Patent, 2:8-10, and is thus encompassed by each of PO’s, Staff’s, and Respondents’ proposed constructions. Additionally, the AAPA “was designed for additive, independent and identically distributed (i.i.d.) noise removal,” and thus is not a deringing filter. ’556 Patent, 2:25-27.

179. In the discussion of claims 1–8 below, I explain the correspondence between Lee and AAPA and each recitation of these claims.

180. One of ordinary skill in the art would have been motivated to combine the teachings of Lee and AAPA for a variety of readily apparent reasons as I explain in more detail below.

A. Claim 1

181. The preamble of claim 1[pre] recites: “**A method comprising:**”

182. The prior art discloses this, to the extent the preamble is limiting, for the same reasons as I explained in paragraph 139.

183. The first limitation of claim 1[a1] recites: “**decoding an image region, wherein the decoding includes performing in-loop filtering to reduce at least a first and a second type of coding artifact using at least two filters in succession,**”

184. The prior art discloses this structure for the same reasons as I explained in paragraphs 140-145.

185. The second limitation of claim 1[a2] recites: “**the at least two filters including a deblocking filter for performing a first pass to reduce blocking artifacts and an adaptive sparse de-noising filter for performing a second pass to reduce noise,**”

186. As I explain below, Lee in combination with the AAPA discloses this limitation under any of the proposed constructions of “sparse de-noising filter.”

187. As I explained in paragraphs 147–150, Lee discloses a deblocking filter 110 performing a first pass of filtering to reduce blocking artifacts. Therefore, after application of Lee’s deblocking filter, the remaining noise would more closely match independent and identically distributed (i.i.d.) noise because blocking artifacts, which are not i.i.d., would have been removed or reduced. ’556 Patent, 14:28-30

(“The signal after deblocking filter is a better match to the independent and identically distributed (i.i.d.) noise presumption.”).

188. A POSITA would have then been motivated to use the AAPA’s adaptive sparse denoising filter, as I explained in paragraphs 117–119, to perform a second pass of filtering to reduce noise because the AAPA adaptive sparse denoising filter offers the advantage of “higher distortion reduction ... than other techniques.” ’556 Patent, 2:43-44. A POSITA would have understood the AAPA adaptive sparse denoising filter to teach a “sparse de-noising filter” under all four proposed constructions because the AAPA “exploits a sparse image model using an over complete set of linear transforms and hard-thresholding,” ’556 Patent, 2:8-10. Additionally, the AAPA “was designed for additive, independent and identically distributed (i.i.d.) noise removal,” and thus is not a deringing filter. *Id.*, 2:25-27.

189. In my opinion, a POSITA would have understood that the main drawback of using the AAPA adaptive sparse denoising filter on its own—namely that it was designed for additive i.i.d. noise such that it “may get confused by true edges or false blocky edges,” and attempting to solve the confusion with “spatio-frequency threshold adaptation ... is not trivial in its implementation,” and “might result into over-smoothed reconstructed pictures or blocky artifact(s)” —would be mitigated or entirely avoided by applying Lee’s deblocking filter in the first pass. ’556 Patent, 2:25-37. Lee’s deblocking filter would remove or greatly reduce blocky

artifacts, and the resulting image, while still having other types of quantization effects not addressed by Lee's deblocking filter, would lack at least the "false blocky edges" that are identified as troublesome for the AAPA adaptive sparse denoising filter. Therefore, it is my opinion that a POSITA would have been motivated to use the AAPA adaptive sparse denoising filter after Lee's deblocking filter because the two would have worked synergistically with each other. Such a combination would have been no more than combining prior art elements (Lee's decoding method using a deblocking filter and second filter, and the AAPA adaptive sparse denoising filter) according to known methods (using a second filter after Lee's deblocking filter) to yield predictable results (improved artifact removal and easier implementation of the AAPA adaptive sparse denoising filter after removal of blocky artifacts). It further would have been no more than applying a known technique (the AAPA adaptive sparse denoising filter) to a known device ready for improvement (Lee's decoder comprising a deblocking filter and at least one other filter) to yield predictable results (overall higher distortion reduction and easier implementation of the AAPA adaptive sparse denoising filter).

190. In my opinion, a POSITA would have further found it obvious to use the AAPA adaptive sparse denoising filter either instead of Lee's deringing filter or in addition to Lee's deringing filter. To the extent that the AAPA adaptive sparse was used in addition to Lee's deringing filter, it is my opinion that a POSITA would

have recognized that it would have been used after Lee's deblocking filter in order to reap the synergy with the Lee's deblocking filter discussed above. In my opinion, the relative order of the AAPA adaptive sparse denoising filter and Lee's deringing filter in such a scenario would have been a design choice that a POSITA would have made based on the demands of the system they were implementing. Thus, a POSITA would have understood that the AAPA adaptive sparse de-noising filter would have performed a "second pass" after Lee's deblocking filter if it was applied between Lee's deblocking and deringing filter. And in any case, in a scenario where both the AAPA adaptive sparse de-noising filter and Lee's deringing filter are used after Lee's deblocking filter, applying the AAPA adaptive sparse de-noising filter before Lee's deringing filter would have been obvious to try as one of two options for ordering these filters.

191. To the extent a POSITA decided to use the AAPA adaptive sparse denoising filter instead of Lee's deringing filter, it is my opinion that doing so would have been no more than a simple substitution of one known element for another to obtain predictable results (overall higher distortion reduction and easier implementation of the AAPA filter). To the extent a POSITA chose to use the AAPA adaptive sparse denoising filter in addition to Lee's deringing filter, it is my opinion that it would have been no more than use of a known technique (the AAPA adaptive sparse denoising filter) to improve similar devices (Lee's decoder using multiple

filters) in the same way (achieving higher overall distortion reduction). In either implementation, a POSITA would have been able to individually tune and optimize filters to reduce artifacts.

192. The third limitation of claim 1[a3] recites: **“wherein the adaptive sparse de-noising filter is selectively enabled or disabled at a given level, the given level being at least a macroblock level, a slice level, a picture level or a sequence level.”**

193. As I explain below, Lee in combination with the AAPA discloses this limitation under any of the four proposed constructions of “sparse de-noising filter.”

194. As I explained in paragraph 152, Lee discloses selectively enabling or disabling its deblocking and deringing filters at a macroblock level. When combining Lee’s filtering method with the AAPA adaptive sparse denoising filter, as I discussed in paragraphs 186–191, it would have been obvious to include a flag for the AAPA adaptive sparse denoising filter because it would allow for the conservation of computing resources by enabling the AAPA adaptive sparse denoising filter only when it is advantageous to apply it. Doing so would have amounted to nothing more than applying a known technique (use of a flag to selectively enable or disable a filter at the macroblock level) to a known method ready for improvement (the AAPA adaptive sparse denoising filter) to yield predictable results (more efficient application of computing resources by deploying

the AAPA adaptive sparse denoising filter only when it is advantageous to do so). Indeed, it was common practice to allow encoders the flexibility of using optional video coding tools, based upon the encoder's access to the original video data and consequent ability to measure whether a tool was effective on a localized basis. *See* Section II.6. For example, flags were commonly used in specifications of adaptive syntax at higher levels because, if a tool is flagged as not being used at lower levels, the syntax related to that tool at lower levels need not appear, thus further reducing overhead data and increasing coding efficiency. *Id.*

B. Claim 2

195. Claim 2 recites: **“The method of claim 1, wherein at least one of the at least two filters is used to enforce at least one compression constraint.”**

196. Lee in combination with the AAPA discloses this limitation under any of the four proposed constructions of “sparse de-noising filter” for the same reasons as I explained in paragraphs 154-157.

C. Claim 3

197. Next claim 3 recites: **“The method of claim 1, wherein at least one of the at least two filters is adaptively enabled or disabled depending upon at least data characteristics corresponding to the image region or data statistics corresponding to the image region.”**

198. As I explain below, Lee in combination with the AAPA discloses this limitation under any of the four proposed constructions of “sparse de-noising filter.”

199. Regardless of the implementation details applied when using a flag to selectively enable or disable the AAPA adaptive sparse denoising filter, as I explained in paragraphs 186–191 and 193–194, the combination of Lee and the AAPA discloses this limitation because Lee’s deblocking filter is also controlled by a “blocking flag indicating whether a need for reducing blocking artifacts near block boundaries exists.” Lee, ¶ 31. For example, Lee teaches how to propagate blocking flag information for inter macroblocks through the use of motion vectors and a reference frame. Lee, ¶¶ 63-65. For example, Lee teaches that its deblocking filtering is not performed when each of the differences between two adjacent blocks and their previous counterparts are not greater than a predetermined value. Lee, ¶ 76. This control of the deblocking filtering is dependent upon data characteristics corresponding to the image region because application of the deblocking filter depends on the differences between blocks in an image region between frames.

D. Claim 4

200. Next claim 4 recites: “**The method of claim 1, wherein a version of the image region that is displayed is different from at least one of an output signal corresponding to the image region after the at least two in-loop filters**

and an initial input signal corresponding to the image region before the at least two in-loop filters.”

201. As I explain below, Lee in combination with the AAPA discloses this limitation under any of the four proposed constructions of “sparse de-noising filter.”

202. Specifically, the reduction of the blocking artifacts by Lee’s deblocking filter and quantization noise by the AAPA adaptive sparse denoising filter results in a “version of the image region that is displayed [being] different from ... an initial input signal corresponding to the image region before the at least two in-loop filters” because the filter change the image to remove the above-noted quantization artifacts.

E. Claim 5

203. Next the preamble of claim 5[pre] recites: **“An apparatus comprising at least a memory and one or more processors, the one or more processors being configured to:”**

204. As I explain below, Lee in combination with the AAPA discloses this limitation under any of the four proposed constructions of “sparse de-noising filter.”

205. Lee discloses that its teachings “may be embodied in a general purpose digital computer that is running a program from a computer usable medium, including but not limited to storage media such as magnetic storage media (e.g., ROM's, floppy disks, hard disks, etc.), optically readable media (e.g., CD-ROMs, DVDs, etc.) and carrier waves (e.g., transmissions over the Internet).” Lee, ¶ 116. It

was well known that a general purpose digital computer comprises one or more processors and that Lee's example of a "computer usable medium" comprises a "memory."

206. The first limitation of the claim, **5[a1]**, recites: **"decode an image region, the one or more processors including at least two filters for successively performing in-loop filtering to reduce at least a first and a second type of coding artifact,"**

207. Lee in combination with the AAPA discloses these limitations for the same reasons as I explained in paragraph 184.

208. The second limitation of the claim, **5[a2]**, recites: **"the at least two filters including a deblocking filter for performing a first pass to reduce blocking artifacts and an adaptive sparse de-noising filter for performing a second pass to reduce noise,"**

209. Lee in combination with the AAPA discloses these limitations for the same reasons as I explained in paragraphs 186–191.

210. The third limitation of the claim, **5[a3]**, recites: **"wherein the adaptive sparse de-noising filter is selectively enabled or disabled at a given level, the given level being at least a macroblock level, a slice level, a picture level or a sequence level."**

211. Lee in combination with the AAPA discloses these limitations for the same reasons as I explained in paragraphs 193–194.

F. Claim 6

212. Claim 6 recites: **“The apparatus of claim 5, wherein at least one of the at least two filters is used to enforce at least one compression constraint.”**

213. Lee in combination with the AAPA discloses this limitation under any of the four proposed constructions of “sparse de-noising filter” for the same reasons as I explained in paragraphs 154-157.

G. Claim 7

214. Next claim 7 recites: **“The apparatus of claim 5, wherein at least one of the at least two filters is adaptively enabled or disabled depending upon at least data characteristics corresponding to the image region or data statistics corresponding to the image region.”**

215. Lee in combination with the AAPA discloses these limitations for the same reasons as I explained in paragraphs 198–199.

H. Claim 8

216. Next claim 8 recites: **“The apparatus of claim 5, wherein a version of the image region that is displayed is different from at least one of an output signal corresponding to the image region after the at least two in-loop filters**

and an initial input signal corresponding to the image region before the at least two in-loop filters.”

217. Lee in combination with the AAPA discloses these limitations for the same reasons as I explained in paragraphs 200–202.

XI. Ground II – Claims 1–8 Are Obvious under 35 U.S.C. § 103 Based on Chiang in View of Lin

218. In my opinion, a person of ordinary skill in the art in 2006 would have found claims 1–8 of the ’556 Patent to have been obvious based on Chiang in view of Lin to the extent that “sparse de-noising filter” is construed as proposed by PO in the related ITC proceeding.

219. One of ordinary skill in the art would have been motivated to combine the teachings of Lin and Chiang for a variety of readily apparent reasons that I explain in more detail below.

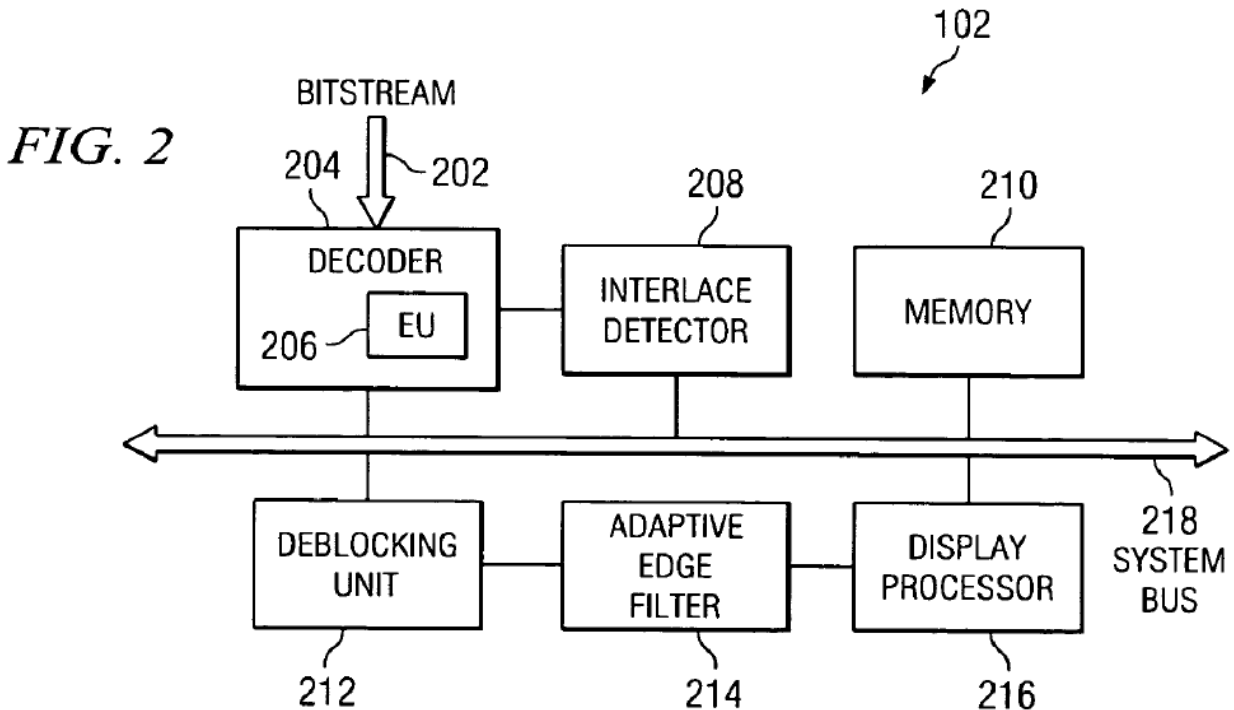
A. Claim 1

220. The preamble of claim 1[pre] recites: “**A method comprising:**”

221. The prior art discloses this, to the extent the preamble is limiting, as follows. Chiang in combination with Lin renders obvious a method. For example, Chiang is directed to “an edge adaptive filtering system for reducing artifacts and method.” Chiang, ¶ 1.

222. The first limitation of claim 1[a1] recites: “**decoding an image region, wherein the decoding includes performing in-loop filtering to reduce at least a first and a second type of coding artifact using at least two filters in succession,**”

223. The prior art discloses this step as follows. Chiang discloses “image decoder 102 [that] decodes [] received information representing an image.” Chiang, ¶ 29. Chiang describes its decoding method with reference to an example image decoder, depicted in Figure 2:



224. Chiang discloses using two types of filters “in conjunction” to remove two types of artifacts from decoded images: an “edge adaptive filtering system that reduces or removes ringing artifacts” and a “de-blocking filtering system that reduces or removes blocking artifacts.” Chiang, ¶ 34.

225. Chiang further teaches that its edge adaptive filtering system and its deblocking filtering system may be implemented as in-loop filters or after the decoding loop. For example, Chiang teaches that its edge adaptive filtering system can be used after motion compensation during image decoding, i.e., during the image decoding loop, ergo the filter is “in-loop.” Chiang ¶ 34 (“The edge adaptive filtering system may be used at any suitable point.... For example [it] could operate after motion compensation during image decoding, after in-loop filtering of decoded images, or before the scaling and processing of the decoded images for display.”), *see also, id.*, ¶ 40 (discussing in-loop deblocking filtering) (“the decoder 204 may implement **in-loop filtering** to reduce or eliminate blocking artifacts”) (my emphasis added). While much of Chiang’s subsequent explanation regarding application of its filters focuses on an embodiment in which the filters are applied outside of the decoding loop, it is my opinion that those teachings equally apply to an in-loop implementation because Chiang’s image decoder is tailorable to its use. In my opinion, it would have been within the skill of a POSITA to reconfigure the filtering process of the DVD/VCD player implementation of Chiang’s modifiable image decoder (Chiang, ¶ 32: “the functional division of FIG. 1 is for illustration only...[its] components...may be combined or omitted and additional components could be added according to particular needs”) for use with a different device (*id.*, “satellite, cable, broadcast, or other television receiver”). A POSITA would have

been motivated to do so, and would have had a reasonable expectation of success, in tuning the filtering process to a particular device to increase the predictability of buffering images. *See, e.g.*, VC1 Standard, 421 (distinguishing between time-delay modes when using, e.g., a DVD as opposed to a video conferencing device).

226. Finally, Chiang teaches that its de-blocking filter and adaptive edge filter are used “in succession.” Specifically, Chiang teaches an implementation in which “[t]he output of the de-blocking unit 212 is provided to an adaptive edge filter 214.” Chiang, ¶ 41. As I explained above, while this particular arrangement involves the de-blocking unit 212 and adaptive edge filter 214 being applied outside of the decoding loop, Chiang teaches applying the same order of filters in-loop.

227. The second limitation of claim 1[a2] recites: **“the at least two filters including a deblocking filter for performing a first pass to reduce blocking artifacts and an adaptive sparse de-noising filter for performing a second pass to reduce noise,”**

228. The prior art discloses this step as follows. As I explained in paragraphs 223–226, Chiang’s deblocking filter “reduces or removes blocking artifacts.” Chiang, ¶ 34. Chiang further teaches that its deblocking filter may be applied before its adaptive edge filter, i.e., in a “first pass.” Chiang, ¶ 41. In this arrangement, Chiang’s adaptive edge filter performs a “second pass” because it filters the image information after it has already been filtered by Chiang’s de-blocking filter.

229. It is also my opinion Chiang's adaptive edge filter is an adaptive "sparse de-noising filter" that reduces noise under PO's current ITC construction. Specifically, Chiang explains that "the adaptive edge filter 214 may perform edge adaptive filtering to reduce or remove ringing artifacts near edges and in flat regions of decoded images." Chiang, ¶¶ 41, 31 ("For each pixel, the edge adaptive filtering system uses the [distance between a pixel of interest and an edge] to identify a filter weight for that pixel"), 81-82 (describing "an inverse relationship between filter weights and relative edge magnitudes"). This is consistent with PO's interpretation of the term, based on its contention that the AV1 CDEF filter is a "sparse de-noising filter" because it "adapts to the sparse nature of underlying video signals by 'preventing excessive blurring when the filter is applied across an edge.'" EX2024, 9. Thus, Chiang's adaptive edge filter is a "sparse de-noising filter" under PO's ITC construction.

230. The third limitation of claim 1[a3] recites: **"wherein the adaptive sparse de-noising filter is selectively enabled or disabled at a given level, the given level being at least a macroblock level, a slice level, a picture level or a sequence level."**

231. The prior art discloses this step as follows. Chiang discloses that the filter weights of its adaptive edge filter are provided by a visual weight map generator. Chiang, ¶ 56 ("adaptive filter 314 also receives the filter weights from the

visual weight map generator 312.”). For each macroblock, the visual weight map generator itself uses visual weight curves output by an edge analyzer, a quality analyzer, a texture analyzer, and a monotone detector to associate filter weights with a given image block. Chiang, ¶¶ 49-50, 52, 55, Fig. 3. Each of these visual weight curves must therefore be individually calculated before filter weights can be provided to the adaptive filter. *See*, Chiang ¶¶ 51-55. However, Chiang discloses circumstances where filter weights are ultimately set to zero, such that no filtering occurs. Chiang, ¶¶ 75, 81, 88. Accordingly, in circumstances where Chiang sets the filter weights to 0 for a given macroblock, Chiang selectively disables its adaptive edge filter (i.e., “adaptive sparse de-noising filter”) at the macroblock level.

232. Chiang also discloses circumstances in which a POSITA would have understood that the filter weight would have been set at or close to zero based on the characteristics of the image or quantization performed on the image. For example, monotone or near monotone regions may be excluded from filtering to “prevent inherent grainy details from being smoothed [out].” Chiang, ¶ 75. As another example, Chiang notes that images that are coarsely quantized may have “high ringing noise due to greater loss of high frequency coefficients” while finely quantized image data may have “little ringing noise and better video quality.” Chiang, ¶ 70.

233. Based on these disclosures, it is my opinion that under certain circumstances, Chiang's adaptive edge filter may perform multiple calculations to ultimately determine that no filtering should occur. Therefore, Chiang's decoder's processing resources may be conserved, and decoding sped up, if the calculations to set filter weights are omitted in circumstances when no filtering is required. In my opinion, a POSITA would have been motivated to look to other art in the field for ways to achieve this.

234. In the same field of invention, Lin discloses a method for decoding that uses in-loop filters. *E.g.*, Lin, ¶¶ 2, 118. Lin explains that various features of loop filtering can be combined. Lin, ¶ 118. For example, Lin discloses that one feature is “[a]daptively enabling or disabling loop filtering in a ... decoder on a per-frame, per-sequence, or other basis.” Lin, ¶ 126. Lin further teaches “[e]nabling or disabling loop filtering in a video decoder according to flags received from a video encoder.” Lin, ¶ 127. Receiving flags from an encoder to enable or disable loop filtering would conserve processing resources by a decoder without degrading image quality in circumstances where the decoder would have otherwise had to perform calculations to determine that no filtering (i.e., setting a filter weight of zero) should occur. Furthermore, Lin teaches using such flags “on a per-frame, per-sequence, or other basis.”

235. Reviewing these teachings, it is my opinion that a POSITA would have been motivated to combine Chiang's decoding method with Lin's teachings of receiving a flag from the encoder to enable or disable loop filtering on a per-frame, per-sequence, or other basis in order to conserve processing resources at the decoder. For example, disabling in-loop filtering as taught by Lin in circumstances when Chiang's filter weights would have been set to zero, such as for decoding monotone or near monotone regions, would have conserved processing resources without degrading image quality. Similarly, in circumstances where an image is very finely quantized, disabling in-loop filtering as taught by Lin would conserve resources while still providing acceptable image quality because finely quantized images already have "little ringing noise and better video quality." Chiang, ¶ 70. As discussed in the background section of the declaration, it was common practice in the video coding field at the relevant time for video coding tools to be enabled and disabled at a variety of levels in the data hierarchy. Encoders in video coding standards had the opportunity to set flags in the bitstream for the tools because the encoder had access to the original (raw) video data, and could evaluate whether or not a tool brings benefit or not at each level of the hierarchy for the current nature of the content.

236. It is also my opinion that a POSITA would have had a reasonable expectation of success in using Lin's flags to enable or disable loop filtering in

combination with Chiang because Chiang already teaches use of a flag to indicate certain image characteristics, such as whether the image is interlaced or progressive to the decoder. Chiang ¶¶ 58, 60, 62. Chiang further teaches that its decoder comprises an “extraction unit 206 [that] extracts various encoding parameters from the bitstream [received from an encoder].” Chiang, ¶ 37. Therefore, Chiang already teaches the use of a flag to control filtering and a way to extract flags from an encoded bitstream. Applying Lin’s teaching of receiving a flag from the encoder to enable or disable filtering on a per-frame, per-sequence, or other basis with Chiang’s decoding method in order to conserve resources by the decoder would have thus been no more than applying known technique (receiving flags from the encoder) to a known method ready for improvement (Chiang’s decoding method) to yield predictable results.

237. Therefore, it is my opinion that Chiang, and/or the combination of Chiang and Lin renders obvious enabling or disabling in-loop filtering, and thus Chiang’s in-loop adaptive edge filter (“adaptive sparse de-noising filter” under the constructions discussed above) at a given level, the given level being at least a sequence level.

B. Claim 2

238. Claim 2 recites: **“The method of claim 1, wherein at least one of the at least two filters is used to enforce at least one compression constraint.”**

239. The combination of Chiang and Lin, which a POSITA would have been motivated to make for the reasons I explained with respect to claim 1, discloses this limitation.

240. First, Chiang discloses its adaptive edge filter adjusting its filtering depending on whether the image being decoded is an interlaced or progressive image “so that interlaced details of the image are protected and filtering is improved or optimized.” Chiang, ¶¶ 31, 37-38. Chiang also discloses its adaptive edge filter enforcing multiple other compression constraints. Chiang explains that encoding parameters that “may include quan_scale_code, quan_scale_type, non_linear_quan_scale, quan_matrix, macroblock_type, and picture_coding_type parameters” are extracted from the MPEG-2 bitstream. *Id.*, ¶ 68. A quality analyzer then uses these parameters to output a visual weigh curve value, ctq. *Id.*, ¶ 71. Chiang then discloses that “edge adaptive filter 214 can be customized with quantization step size to select suitable visual weight curve values for balancing between filtering ringing noise and smoothing detail.” Chiang, ¶ 70, *see also, id.*, ¶¶ 71-74. In all cases, Chiang’s adaptive edge filter changes its filtering based on the qualities of the compressed image being filtered. The adaptive edge filter therefore enforces a compression constraint because it alters its filtering, in a predetermined manner known to both the encoder and decoder based on how the image is encoded (interlaced vs. progressive) and the various encoding parameters noted above.

241. Lin discloses using a flag to enable or disable loop filtering, and a POSITA would have been motivated to incorporate this teaching from Lin with Chiang as I discussed with regard to claim 1. Enabling or disabling loop filtering via a flag, as taught by Lin, would result in both of Chiang's deblocking and adaptive edge filters enforcing a compression constraint because both the encoder and decoder are constrained ahead of time to use syntax and tools available to ensure that which is encoded can later be decoded.

242. Thus, the combination of Chiang and Lin teaches multiple ways in which Chiang's adaptive edge filter (and deblocking filter, in the case of the use of flags) enforce a compression constraint.

C. Claim 3

243. Next claim 3 recites: **“The method of claim 1, wherein at least one of the at least two filters is adaptively enabled or disabled depending upon at least data characteristics corresponding to the image region or data statistics corresponding to the image region.”**

244. The prior art discloses this step as follows. As I explained in paragraphs 231–237, the combination of Chiang and Lin renders obvious enabling or disabling in-loop filtering by using a flag from the encoder in circumstances where filter weights would have been set to zero or where images are finely quantized and thus are already relatively high quality. Therefore both (i.e., “at least one”) of Chiang's

in-loop filters are adaptively enabled or disabled depending on data characteristics corresponding to the image region (e.g., whether the image region is monotone or nearly monotone, Chiang, ¶ 75, or finely quantized, Chiang, ¶ 70).

D. Claim 4

245. Next claim 4 recites: **“The method of claim 1, wherein a version of the image region that is displayed is different from at least one of an output signal corresponding to the image region after the at least two in-loop filters and an initial input signal corresponding to the image region before the at least two in-loop filters.”**

246. The prior art discloses this step as follows. Chiang explains that its deblocking filter “at least partially reduce[s] blocking artifacts in the image” and its adaptive edge filter “partially reduce[s] ringing artifacts in the image.” Chiang, ¶ 29. Therefore, the reduction of the blocking and ringing artifacts by Chiang’s deblocking and adaptive edge filters results in a “version of the image region that is displayed [being] different from ... an initial input signal corresponding to the image region before the at least two in-loop filters.”

E. Claim 5

247. Next the preamble of claim 5[pre] recites: **“An apparatus comprising at least a memory and one or more processors, the one or more processors being configured to:”**

248. To the extent it is limiting, the prior art discloses this structure as follows. Chiang discloses an apparatus comprising an image decoder. Chiang, ¶ 27, Fig. 1. Chiang further discloses that the image decoder comprises a display processor 216 and a memory 210. Chiang, ¶¶ 42-43, Fig. 2. Further, it is clear that in addition to the processor and memory explicitly identified in Figure 2, the other elements of the decoding apparatus disclosed by Chiang would have also been implemented via one or more processors and memories.

249. The first limitation of the claim, 5[a1], recites: **“decode an image region, the one or more processors including at least two filters for successively performing in-loop filtering to reduce at least a first and a second type of coding artifact,”**

250. Chiang in combination with Lin renders obvious this limitation for the same reasons as I explained in paragraphs 223–226.

251. The second limitation of the claim, 5[a2], recites: **“the at least two filters including a deblocking filter for performing a first pass to reduce blocking artifacts and an adaptive sparse de-noising filter for performing a second pass to reduce noise,”**

252. Chiang in combination with Lin renders obvious this limitation for the same reasons as I explained in paragraphs 227-229.

253. The third limitation of the claim, **5[a3]**, recites: “**wherein the adaptive sparse de-noising filter is selectively enabled or disabled at a given level, the given level being at least a macroblock level, a slice level, a picture level or a sequence level.**”

254. Chiang in combination with Lin renders obvious this limitation for the same reasons as I explained in paragraphs 231–237.

F. Claim 6

255. Claim 6 recites: “**The apparatus of claim 5, wherein at least one of the at least two filters is used to enforce at least one compression constraint.**”

256. Chiang in combination with Lin discloses this limitation for the same reasons as I explained in paragraphs 239-242.

G. Claim 7

257. Next claim 7 recites: “**The apparatus of claim 5, wherein at least one of the at least two filters is adaptively enabled or disabled depending upon at least data characteristics corresponding to the image region or data statistics corresponding to the image region.**”

258. Chiang in combination with Lin renders obvious this limitation for the same reasons as I explained in paragraph 243.

H. Claim 8

259. Next claim 8 recites: **“The apparatus of claim 5, wherein a version of the image region that is displayed is different from at least one of an output signal corresponding to the image region after the at least two in-loop filters and an initial input signal corresponding to the image region before the at least two in-loop filters.”**

260. Chiang in combination with Lin renders obvious this limitation for the same reasons as I explained in paragraphs 245–246.

XII. Ground III – Claims 1–8 Are Obvious under 35 U.S.C. § 103 Based on VC1 Standard in View of Guleryuz

261. In my opinion, a person of ordinary skill in the art in 2006 would have found claims 1–8 of the ’556 Patent to have been obvious based on the VC1 Standard in view of Guleryuz to the extent that “sparse de-noising filter” is construed under any of the four proposed constructions in the related ITC proceeding.

262. In the discussion of claims 1–8 below, I explain the correspondence between the VC1 Standard and Guleryuz and each recitation of these claims.

263. One of ordinary skill in the art would have been motivated to combine the teachings of VC1 and Guleryuz for a variety of readily apparent reasons that I describe in more detail below.

A. Claim 1

264. The preamble of claim 1[pre] recites: **“A method comprising:”**

265. The prior art discloses this step, to the extent the preamble is limiting, as follows. The VC1 Standard in combination with Guleryuz discloses or renders obvious a method. For example, the VC1 Standard discloses a decoding process. *See, e.g.*, VC1, 4-5; Figs. 3-4.

266. The first limitation of claim 1[a1] recites: **“decoding an image region, wherein the decoding includes performing in-loop filtering to reduce at least a first and a second type of coding artifact using at least two filters in succession,”**

267. The prior art discloses this step as follows. The VC1 Standard discloses a decoding process that includes an in-loop deblock filtering process. VC1, 178-188 and 322-328. *See generally*, VC1, Section 8 and Section 10. The in-loop deblock filtering process reduces a first type of coding artifact – blocking artifacts. VC1, 178 (“the intent of loop filtering is to smooth out the discontinuities at block boundaries”).

268. In the same field as the VC1 Standard, Guleryuz discloses a “high-performance, nonlinear, loop filter that reduces quantization noise over video frames composed of locally uniform regions (smooth, high frequency, texture, etc.) separated by singularities.” Guleryuz, Abstract. Guleryuz explains that its filter “is not limited to block based transform coders and provides a robust solution for intra as well as differentially encoded frames.” *Id.* Guleryuz’s filter reduces quantization noise. Guleryuz, 1-2. Such noise includes both blocking artifacts and artifacts around

sharp edges. Guleryuz, 3-4. Therefore, Guleryuz's filter reduces at least a second type of coding artifact (e.g., artifacts around sharp edges).

269. A POSITA looking to apply the teachings of the VC1 Standard would have also recognized that in-loop deblocking filters such as the one taught by the VC1 Standard have shortcomings in that they do not address other types of coding artifacts very well. The stated purpose of the VC1 in-loop deblocking filter is to "smooth out the discontinuities at block boundaries" – it is silent regarding other coding artifacts. VC1, 798. Therefore, using a second in-loop filter to remove coding artifacts after application of a deblocking filter was a common technique to improve the quality of decoded images. *See, e.g.,* Lee, ¶¶ 21, 78, 98 (disclosing first and second "filterings," where the second is a "weak deblocking filtering" intended to change fewer pixel values); Chiang, ¶ 34 (describing the reconfigurable, and combinatorial, application of edge adaptive and de-blocking filtering to remove image ringing and blocking artifacts). Indeed, the VC1 Standard already teaches the use of multiple in-loop filters, each addressing different coding artifacts, because it employs both an overlap smoothing filter and a deblocking filter in the decoding loop. E.g., VC1 at Figs. 3, 4. And while the VC1 Standard does disclose use of a deblocking and deringing filter outside of the coding loop, these filters do not present the benefit of providing a higher quality reference frame that in-loop filtering provides. This benefit of in-loop filtering was well understood in the art. E.g.,

Guleryuz, 4 (“We expect our results to improve with the incorporation of more sophisticated processing.”) (discussing shortcomings of results due to application of the filter in post-processing rather than in-loop). Therefore, it is my opinion that a POSITA would have been motivated to combine the teachings of the VC1 Standard and Guleryuz by applying Guleryuz’s in-loop denoising filter after the VC1 Standard in-loop deblock filter because doing so would have amounted to no more than combining prior art elements (the VC1 Standard in-loop deblock filter and Guleryuz’s in-loop denoising filter) according to known methods (implementing an in-loop denoising filter after an in-loop deblocking filter as taught in numerous prior art references including Lee and Chiang) to yield predictable results (improved decoded image quality over use of just VC1’s in-loop deblock filter). Likewise, it would have been no more than use of a known technique (applying an in-loop denoising filter after an in-loop deblocking filter, as taught by numerous prior art references such as Lee and Chiang) to a known method ready for improvement (the VC1 Standard applying an in-loop deblock filter) to yield predictable results (improved decoded image quality).

270. In combining the VC1 Standard in-loop deblock filter and Guleryuz’s in-loop denoising filter, it is my opinion that a POSITA would have been motivated to apply Guleryuz’s in-loop denoising filter after the VC1 Standard in-loop deblock filter. In particular, Guleryuz notes the fact that quantization noise is not i.i.d. noise

presents an obstacle to applying its denoising filter. Guleryuz, 2. Therefore after removal or reduction of blocking artifacts, the remaining quantization noise is a much closer match to i.i.d. noise, as the '556 patent admits was known. *See, e.g.,* AAPA, '556 Patent, 2:25-27. Accordingly, a POSITA would have been motivated to apply Guleryuz's in-loop denoising filter after the VC1 Standard in-loop deblock filter because doing so would have made implementing Guleryuz's filter easier. Further, the application of a denoising filter (such as that taught by Guleryuz) after a deblocking filter (such as that taught by the VC1 Standard) was a well-known technique for improving image quality due to the synergistic properties of these two types of filters. *See, e.g.,* Lee at Fig. 1, Chiang at Fig. 2, EX1015, section 2. However, even if a POSITA would not have made the connection that applying the deblock filter first would have eased implementation of Guleryuz's denoising filter, doing so would have been obvious to try because it is one of two options (deblock filter first or denoising filter first). It also would have been obvious to try because other prior art such as Lee and Chiang teach applying a deblocking filter first. *E.g.,* Lee, Fig. 1; Chiang, ¶ 41.

271. Therefore, it is my opinion that the combination of the VC1 Standard and Guleryuz discloses decoding an image region wherein the decoding includes performing in-loop filtering to reduce at least a first (block artifact) and a second

(quantization noise) type of coding artifact using at least two filters (VC1 in-loop deblock filter and Guleryuz's in-loop denoising filter) in succession.

272. The second limitation of claim 1[a2] recites: “**the at least two filters including a deblocking filter for performing a first pass to reduce blocking artifacts and an adaptive sparse de-noising filter for performing a second pass to reduce noise,**”

273. The prior art discloses this step as follows. As I explained in paragraphs 267–271, a POSITA would have been motivated to combine the teachings of the VC1 Standard and Guleryuz such that the VC1 Standard's in-loop deblock filter is applied first to reducing blocking artifacts (i.e., in a “first pass”). As such, Guleryuz's in-loop denoising filter would have been applied in a “second pass.” And as I explained in paragraphs 267–271, Guleryuz's in-loop denoising filter reduces quantization noise. Guleryuz, 1.

274. Guleryuz's filter teaches a “sparse denoising filter” under any of the four proposed constructions in the related ITC proceeding. Indeed, Guleryuz describes its “sparse denoising filter” using almost identical language to the Respondents' alternative construction for the term (and consequently the '556 Patent's description of the term). *Compare* Guleryuz, 1 (“a **sparse image model** using a possibly **overcomplete set of linear transforms.**”) (my emphases); *with* '556 Patent, 2:7-10 (“a nonlinear de-noising filter adapts to non-stationary image

statistics which exploits **a sparse image model** using an **over complete set of linear transforms** and hard-thresholding.”) (my emphases), 2-3 (“...we **apply a hard-thresholding** nonlinearity using a threshold to preprocess...” (my emphases); ’556 Patent, 13:22-25 (“an adaptive sparse de-noising filter operation involves the following steps...**an over complete transform**”) (my emphases). Further, Guleryuz’s filter is also consistent with the Staff’s construction because Guleryuz’s transforms would have resulted in a “sparse representation of the signal” as in the Staff’s construction. And because PO’s construction broadly encompasses both sparse signals and sparse representations of signals, Guleryuz’s filter is also consistent with PO’s construction. Finally, Guleryuz’s in-loop denoising filter functions as a “sparse denoising filter” even if the Patent Owner disclaimed deringing filters during prosecution because Guleryuz’s filter reduces all quantization noise, not just ringing artifacts.

275. Guleryuz’s filter is also adaptive. Specifically, Guleryuz explains that “[c]onsideration is given to compression parameters/modes to ensure overfiltering is kept to a minimum and a refinement step is proposed that improves the expectation calculation.” Guleryuz, 2. Guleryuz further teaches its filter calculating multiple denoised estimates of a signal and averaging them “to form an overall estimate” used to set filter weights. Guleryuz, 2-3. Therefore, Guleryuz’s sparse denoising filter is

adaptive because filter weights are changed (i.e., “adapted”) based on changes in the signal. Guleryuz, 3.

276. The third limitation of claim 1[a3] recites: **“wherein the adaptive sparse de-noising filter is selectively enabled or disabled at a given level, the given level being at least a macroblock level, a slice level, a picture level or a sequence level.”**

277. The prior art discloses this step as follows. The VC1 Standard teaches use of a loop filter flag that is “a Boolean that shall indicate whether loop filtering is enabled for the sequence.” VC1, 462-63. The VC1 Standard teaches that if “LOOPFILTER == 0, then loop filtering shall not be enabled. If LOOPFILTER == 1, then loop filtering shall be enabled.” *Id.*

278. In the combination I explained in paragraphs 267–271, where Guleryuz’s adaptive sparse de-noising in-loop filter is used after the VC1 Standard’s deblock filter, it is clear that the VC1 Standard’s loop filter flag would selectively enable or disable Guleryuz’s adaptive sparse de-noising filter at the sequence level because Guleryuz’s filter is applied in-loop (after the VC1 Standard deblock filter). Indeed, the VC1 Standard loop filter flag would have applied to both the in-loop deblock filter taught by the VC1 and Guleryuz’s adaptive sparse de-noising filter because the two filters would both be in the same “loop filter” (red) function block I annotate in VC1, Figure 3:

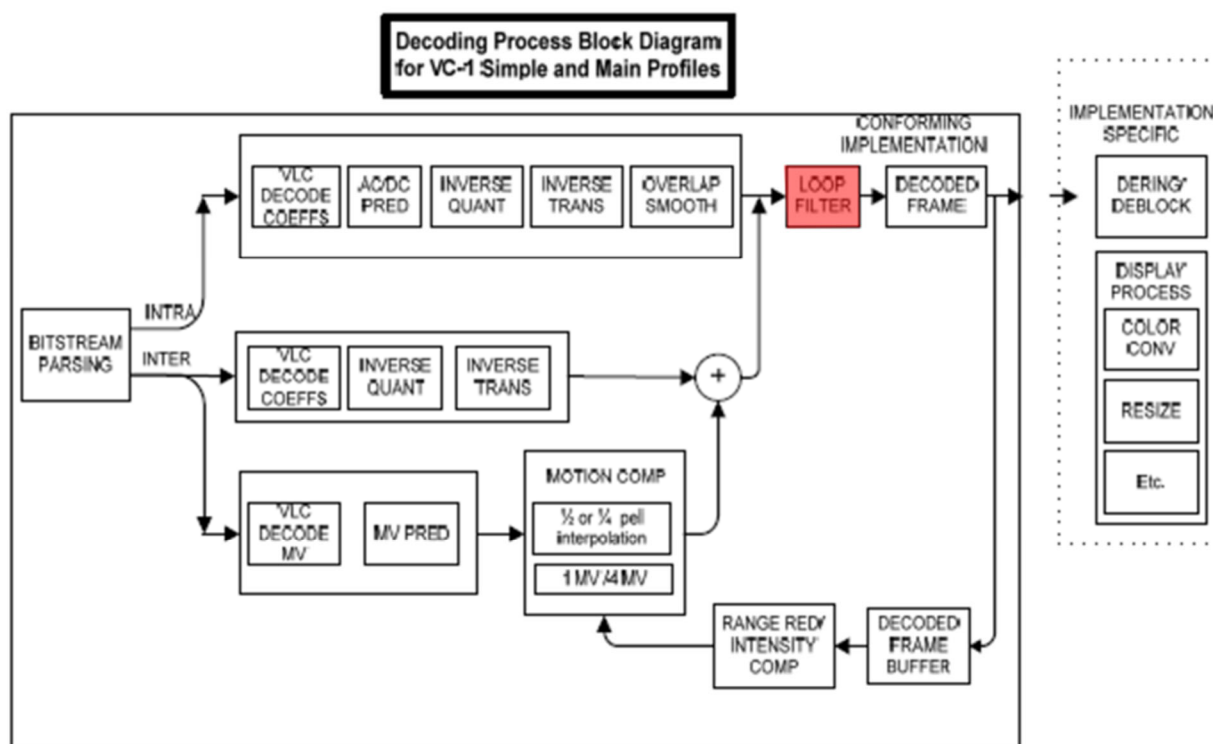


Figure 3: Decoding Process Block Diagram for Simple and Main Profile

279. I have been asked to assume that PO will argue that the VC1 Standard's LOOPFILTER flag would not have selectively enabled or disabled Guleryuz's adaptive sparse de-noising filter in the combination as proposed above. PO would be incorrect because a POSITA would have implemented Guleryuz's adaptive sparse denoising filter in the same function block as VC1's in-loop deblocking filter, such that the LOOPFILTER flag would selectively enable or disable both filters together. In my opinion, a POSITA would have found it obvious to use a single flag to control two filters and would have been motivated to do so because the VC1 Standard itself

teaches that a single flag, the POSTPROC flag, can control two separate filters. VC1 at 472.

280. Alternatively, and as a design choice that would offer finer control over in-loop filtering at the expense of slightly reduced compression due to having to code an additional filter flag, a POSITA would have also been motivated to implement Guleryuz's adaptive sparse de-noising filter in the combination with its own flag that selectively enables or disables the filter. The VC1 Standard itself teaches use of flags to selectively enable or disable various filters in the decoding processes. In addition to the LOOPFILTER flag, the VC1 Standard also teaches an OVERLAP flag to selectively enable or disable a filtering operation referred to as overlap smoothing, VC1, 195, and POSTPROC field to selectively enable or disable post-processing filtering of pictures, VC1, 472. Therefore, a POSITA would have been motivated to individually control Guleryuz's adaptive sparse de-noising filter as applied after VC1's in-loop deblocking filter with a separate flag. Doing so would have been no more than combining prior art elements (the VC1 Standard's decoding process and Guleryuz's adaptive sparse de-noising filter) according to known methods (VC1's use of flags to enable or disable filters used in the decoding process) to yield predictable results (selectively enabling or disabling Guleryuz's adaptive sparse de-nosing filter). It also would have merely amounted to use of a known technique (VC1's use of flags to enable or disable filters in the decoding process) to

improve similar methods in the same way (also using a flag to enable or disable Guleryuz's adaptive sparse de-noising filter when it is applied in VC1's decoding process).

281. Therefore, it is my opinion that the combination of the VC1 Standard and Guleryuz teaches at least two possible configurations that a POSITA would have been motivated to implement in which Guleryuz's adaptive sparse denoising filter is selectively enabled or disabled at the sequence level.

B. Claim 2

282. Claim 2 recites: **"The method of claim 1, wherein at least one of the at least two filters is used to enforce at least one compression constraint."**

283. The combination of the VC1 Standard and Guleryuz discloses this limitation.

284. Guleryuz teaches that its adaptive sparse denoising filter includes a "mask function" used to reduce overfiltering in circumstances where differentially coded SKIP macroblocks have been previously filtered. Guleryuz, 3. Specifically, Guleryuz teaches that the mask function "determines the pixels which should not undergo filtering." Guleryuz, 3. "In short, if the mask for a pixel is set to 1, the final filtered value is copied from the output of the *combine* function, otherwise the final filtered value is set to the value provided by *y*." *Id.* (emphasis in original). "For intra frames the mask function is 1 on all pixels." *Id.* The value of the mask for a pixel is

determined based on “the compression modes around each macroblock boundary.”

Id. In other words, if a frame is intra coded, its pixels do not undergo filtering by Guleryuz’s filter. Otherwise, the mask function determines the strength of filtering as disclosed by Guleryuz.

285. Guleryuz provides rules for determining the mask value for six different macroblock types. *Id.*, section 3.5. The rules are copied below:

The coefficient based thresholds are determined as follows:

if $h_{i,j}$ overlaps an INTRAQ or a PQ or a PQM MB
the threshold remains T ,

else if $h_{i,j}$ overlaps a PQ1 MB
the threshold becomes $7/8T$,

else if $h_{i,j}$ overlaps a SKIPM or a SKIP MB,
the threshold becomes $1/2T$,

else
 $h_{i,j}$ passes the initial check and $E[c_i(j)|\Psi] = d_i(j)$.

Assuming $B \times B$ macroblocks, the conditional used for determining I in the reported simulations is given by:

if a boundary is next to an INTRAQ or a PQ or a PQM MB,
 $I = B/2$,

else if a boundary is next to a PQ1 or a SKIPM or a SKIP MB,
 $I = B/4$,

else
 $I = 0$.

Guleryuz’s adaptive sparse denoising filter thus enforces a compression constraint because it values the strength of filtering (or entirely disables filtering) based on

compression characteristics of the macroblock being filtered (either, what type of macroblock it is coded as).

C. Claim 3

286. Next claim 3 recites: **“The method of claim 1, wherein at least one of the at least two filters is adaptively enabled or disabled depending upon at least data characteristics corresponding to the image region or data statistics corresponding to the image region.”**

287. The prior art discloses this step as follows. Guleryuz teaches that its adaptive sparse denoising filter includes a “mask function” used to reduce overfiltering in circumstances where differentially coded SKIP macroblocks have been previously filtered. Guleryuz, 3. Specifically, Guleryuz teaches that the mask function “determines the pixels which should not undergo filtering.” Guleryuz, 3. “In short, if the mask for a pixel is set to 1, the final filtered value is copied from the output of the combine function, otherwise the final filtered value is set to the value provided by y.” *Id.* Therefore, Guleryuz teaches adaptively disabling (via a mask value of 1) or enabling (via a mask value other than 1) its adaptive sparse denoising filter.

288. The value of the mask for a pixel is determined based on “the compression modes around each macroblock boundary.” Guleryuz, 3. Guleryuz thus adaptively determines the mask value depending on “data characteristics

corresponding to the image region,” i.e., the compression modes around each macroblock. Therefore, it is my opinion that the combination of the VC1 Standard and Guleryuz teaches adaptively enabling or disabling the adaptive sparse denoising filter as taught by Guleryuz “depending on data characteristics corresponding to the image region.”

D. Claim 4

289. Next claim 4 recites: **“The method of claim 1, wherein a version of the image region that is displayed is different from at least one of an output signal corresponding to the image region after the at least two in-loop filters and an initial input signal corresponding to the image region before the at least two in-loop filters.”**

290. The prior art discloses this step as follows. The VC1 Standard teaches that its in-loop deblock filter “smooth[es] out the discontinuities at block boundaries.” VC1, 178. Similarly, Guleryuz teaches that its adaptive sparse denoising filter produces “filter effects [that] are especially pronounced around sharp edges.... texture and edge information is preserved while artifacts are suppressed.” Guleryuz, 3. Based on each of these teachings, because each of the VC1 Standard deblock filter and Guleryuz’s adaptive sparse denoising filter modifies an input image to produce a filtered output image, the version of the image that is displayed after the in-loop filtering performed by the combination of the VC1 Standard and

Guleryuz I explained in paragraphs 267–271 would be different from the initial input signal corresponding to the image before the two in-loop filters are applied.

E. Claim 5

291. Next the preamble of claim 5[pre] recites: “**An apparatus comprising at least a memory and one or more processors, the one or more processors being configured to:**”

292. To the extent it is limiting, the prior art discloses this structure as follows. Clearly both the VC1 Standard and the adaptive sparse denoising filter taught by Guleryuz are intended to be implemented by a decoder, and that decoders are apparatuses comprising at least a processor and memory because the VC1 Standard acknowledges that “processing speed and memory” are taken into consideration by the standard. VC1, 449.

293. The first limitation of the claim, 5[a1], recites: “**decode an image region, the one or more processors including at least two filters for successively performing in-loop filtering to reduce at least a first and a second type of coding artifact,**”

294. The VC1 Standard in combination with Guleryuz discloses or renders obvious this limitation for the same reasons as I explained in paragraphs 267–271.

295. The second limitation of the claim, 5[a2], recites: “**the at least two filters including a deblocking filter for performing a first pass to reduce**

blocking artifacts and an adaptive sparse de-noising filter for performing a second pass to reduce noise,”

296. The VC1 Standard in combination with Guleryuz discloses or renders obvious this limitation for the same reasons as I explained in paragraphs 273–275.

297. The third limitation of the claim, **5[a3]**, recites: **“wherein the adaptive sparse de-noising filter is selectively enabled or disabled at a given level, the given level being at least a macroblock level, a slice level, a picture level or a sequence level.”**

298. The VC1 Standard in combination with Guleryuz discloses or renders obvious this limitation for the same reasons as I explained in paragraphs 277–281.

F. Claim 6

299. Claim 6 recites: **“The apparatus of claim 5, wherein at least one of the at least two filters is used to enforce at least one compression constraint.”**

300. The VC1 Standard in combination with Guleryuz discloses this limitation for the same reasons as I explained in paragraphs 283–285.

G. Claim 7

301. Next claim 7 recites: **“The apparatus of claim 5, wherein at least one of the at least two filters is adaptively enabled or disabled depending upon at least data characteristics corresponding to the image region or data statistics corresponding to the image region.”**

302. The VC1 Standard in combination with Guleryuz discloses or renders obvious this limitation for the same reasons as I explained in paragraphs 287–288.

H. Claim 8

303. Next claim 8 recites: **“The apparatus of claim 5, wherein a version of the image region that is displayed is different from at least one of an output signal corresponding to the image region after the at least two in-loop filters and an initial input signal corresponding to the image region before the at least two in-loop filters.”**

304. The VC1 Standard in combination with Guleryuz discloses or renders obvious this limitation for the same reasons as I explained in paragraphs 289–290.

XIII. Summary

305. It is my opinion that Grounds IA, IB, 2, and 3 render Challenged Claims 1–8 unpatentable as obvious over the Lee, AAPA, Chiang, Lin, VC1 Standard, and Guleryuz prior art references as I explained in the paragraphs preceding.

XIV. Additional Remarks

306. 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.