

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

GOOGLE LLC
Petitioner

v.

UNILOC 2017 LLC,
Patent Owner

Patent No. 6,519,005

**DECLARATION OF MAJID RABBANI, PH.D.
IN SUPPORT OF PETITION FOR *INTER PARTES* REVIEW
OF U.S. PATENT 6,519,005**

I.	INTRODUCTION	1
II.	BACKGROUND AND QUALIFICATIONS	2
III.	MATERIALS CONSIDERED AND SUMMARY OF OPINIONS.....	6
IV.	PERSON OF ORDINARY SKILL IN THE ART.....	8
V.	TECHNICAL BACKGROUND	10
A.	Overview of Video Compression/Video Encoding	11
B.	Pixel Arrays and Macroblocks.....	15
C.	Motion Estimation	16
D.	Block Matching.....	20
E.	Prediction Modes and Selection of an Optimum Prediction Mode	23
F.	Concurrent/Parallel Performance of Motion Estimation In Multiple Prediction Modes.....	30
VI.	THE '005 PATENT	38
A.	Summary of the '005 Patent.....	38
B.	Prosecution History of the '005 Patent.....	42
VII.	CLAIM CONSTRUCTION	44
VIII.	OVERVIEW OF THE PRIOR ART.....	47
A.	<i>Hanami</i> (Ex. 1005)	47
1.	<i>Hanami</i> 's Overall Structure and Method.....	51
2.	Template Block In Current Frame and Search Area In Reference Frame.....	54
3.	Block-matching, Evaluation Values, and Motion Vector Generation	56
4.	Prediction Modes.....	59
B.	<i>Mitchell</i> (Ex. 1006).....	62

IX.	THE PRIOR ART DISCLOSES ALL OF THE FEATURES OF CLAIMs 1-3, 5, 6, 8, 39 and 41 OF THE '005 PATENT.....	66
A.	<i>Hanami</i> Discloses the Features of Claims 1-3, 5, 6, 8, 39 and 41	66
1.	Claim 1	66
a)	“A method for motion coding an uncompressed digital video data stream, including the steps of:”.....	66
b)	[1] “comparing pixels of a first pixel array in a picture currently being coded with pixels of a plurality of second pixel arrays in at least one reference picture and”.....	69
	[2] “concurrently performing motion estimation for each of a plurality of different prediction modes in order to determine which of the prediction modes is an optimum prediction mode;”.....	69
1)	“ <i>Concurrently</i> ” Under First Possible Construction....	84
2)	“ <i>Optimum</i> ” Under First and Second Possible Constructions.....	86
c)	“determining which of the second pixel arrays constitutes a best match with respect to the first pixel array for the optimum prediction mode; and,”....	88
1)	“ <i>Best match</i> ” Under First and Second Possible Constructions.....	93
d)	“generating a motion vector for the first pixel array in response to the determining step.”.....	95
2.	Claim 2	100
a)	“The method as set forth in claim 1, wherein the first and second pixel arrays each have a size and structure defined by an MPEG standard.”.....	100
3.	Claim 3	107

a)	“The method as set forth in claim 1, wherein the method is implemented using a motion estimation search engine of a digital video encoder.”	107
4.	Claim 5	110
a)	“The method as set forth in claim 1, wherein the motion coding is performed in accordance with an MPEG standard.”	110
5.	Claim 6	113
a)	“The method as set forth in claim 5, further comprising the initial step of providing information identifying a picture type of the first pixel array and using this information in the comparing step.”	113
6.	Claim 8	115
a)	“The method as set forth in claim 1, wherein the different prediction modes are: frame prediction mode for frame pictures; and, field prediction mode for frame pictures.”	115
7.	Claim 39.....	117
a)	“A motion estimation search system that concurrently performs motion estimation using each of a plurality of different motion estimation prediction modes and then selects the prediction mode that produces the optimum result.”	117
8.	Claim 41.....	117
a)	“A method, including the steps of:”	117
b)	“concurrently performing motion estimation using each of a plurality of different motion estimation prediction modes;”	118
c)	“comparing the results produced using each different prediction mode; and,”	118
d)	“selecting the prediction mode that produced an optimum result; and,”	118

e)	“generating one or more motion vectors using the selected prediction mode.”	118
B.	<i>Hanami</i> and <i>Mitchell</i> Disclose Claims 2 and 5	118
1.	Claim 2	118
2.	Claim 5	123
X.	CONCLUSION	139

I, Majid Rabbani, declare as follows:

I. INTRODUCTION

1. I have been retained by Google LLC (“Petitioner”) as an independent expert consultant in this proceeding before the United States Patent and Trademark Office (“PTO”) regarding U.S. Patent No. 6,519,005 (“the ’005 patent”) (Ex. 1001)¹. I have been asked to consider whether certain references disclose or suggest the features recited in claims 1-3, 5, 6, 8, 39 and 41 (“the challenged claims”) of the ’005 patent. My opinions are set forth below.

2. I am being compensated at my rate of \$450 per hour for the time I spend on this matter. My compensation is in no way contingent on the nature of my findings, the presentation of my findings in testimony, or the outcome of this or any other proceeding. I have no other interest in this proceeding.

¹ Where appropriate, I refer to exhibits attached to the petition for *Inter Partes* Review of the ’005 patent.

II. BACKGROUND AND QUALIFICATIONS

3. Presently, I am a visiting full professor at the Rochester Institute of Technology (RIT) in the Department of Electrical Engineering, where I have held the position of Adjunct professor of Electrical Engineering in the Department of Electrical Engineering (1985-1992 and 2011-2016). I mostly teach graduate level courses through the Department of Electrical Engineering at the Rochester Institute of Technology. The courses I have taught include Digital Image Processing (EEEE779), Digital Signal Processing (EEEE678), Image and Video Compression (EEEE781), Information Theory (EEEE794), Circuits I (EEEE281) and Pattern Recognition (EEEE670). Image and video compression technology is the part of the material that I have covered in Information Theory, Digital Image Processing and Image and Video Compression courses. I understand that a copy of my current curriculum vitae is attached as Exhibit 1003.

4. My formal education includes a bachelor's degree in Electrical Engineering from the AryaMehr (now Sharif) Institute of Technology, Tehran, Iran in 1978, a master's degree in Electrical and Computer Engineering from the University of Wisconsin-Madison in 1979, and a Ph.D. degree in Electrical and Computer Engineering from the University of Wisconsin-Madison in 1983.

5. After completing my Ph.D in 1983, I joined the Kodak research laboratories in Rochester, NY, and in 2016 I retired with the rank of *Kodak Fellow*. While at Kodak, I initiated research and built the competencies and teams in three diverse technology areas new to Kodak, including, image compression, video processing and analysis, and image security and watermarking.

6. In the 25-year period spanning 1988-2012, I assumed multiple leadership roles in representing Kodak at world-renowned and internationally prominent standardization organizations such as the (i) Joint Photographic Experts Group (JPEG), which is the joint committee between ISO/IEC JTC 1² and ITU-T³ that created and maintains the JPEG, JPEG 2000, and JPEG XR standards, and (ii) the Moving Pictures Expert Group (MPEG), which also formed under ISO and ITU-

² ISO is the International Organization for Standardization and IEC is the International Electrotechnical Commission. ISO/IEC JTC 1 is a Joint Technical committee of ISO and IEC for the purpose of developing, maintaining, and promoting standards in the field of Information Technology (IT) and Information and Communication Technology (ICT).

³ ITU is the International Telecommunication Union and ITU-T is the Telecommunication Standardization Sector of ITU-T.

T⁴. Further, during 1998-2000, as part of my leadership activities in the JPEG group, I was the Chairman of the *Coding Efficiency Group* which included more than 50 scientists of the JPEG-2000 committee. I was also elected by the Digital Cinema Initiatives (DCI) Consortium as the industry expert to the *Digital Cinema Compression Experts Panel* during 2003-2004.

7. In the past 30 years, I have taught about 50 short courses and seminars at the various *Society of Photo-Optical Instrumentation Engineers* (SPIE), Society for Imaging Science and Technology (IS&T) and Institute of Electrical and Electronics Engineers (IEEE) conferences. In January 2012, I was awarded the “Electronic Imaging Distinguished Educator Award,” by IS&T and SPIE for teaching short courses in the various areas of electronic imaging at the SPIE/IS&T flagship conference, the EI Symposium, for 25 consecutive years in the period of 1988-2012. Since then, I have taught for seven more consecutive years at the same conference.

⁴ MPEG’s official designation is ISO/IEC JTC 1/Sc 29/WG 11. SC29 is a subcommittee under JTC 1 for coding of audio, picture, multimedia and hypermedia information for coding of audio, picture, multimedia and hypermedia information. SC29 oversees Working Group 11 (WG11), also known as MPEG.

8. I am a *Fellow* of IEEE (1997), a *Fellow* of SPIE (1993), a *Kodak Fellow* and a *Kodak Distinguished Inventor*. I was interviewed by the *El Pais* Spanish newspaper in 2003 regarding my leadership role in JPEG-2000; by the Rochester *Democrat and Chronicle* newspaper in 2001 for my contributions to digital watermarking used in digital cinema; and by the *Persian Heritage* magazine for my overall technical accomplishments.

9. My research interest span the various aspect of digital image and video processing and analysis where I have delivered 40+ keynote and plenary presentations, have published over 60 technical articles, four books chapters, and hold 44 issued US patents. I am the principle author of the book *Digital Image Compression Techniques* published by SPIE Press in 1991 and the Editor of the SPIE Milestones Series on *Image Coding and Compression*, published in 1992. A comprehensive list of my patents, conference and journal publications, and keynote and plenary speeches is included in the copy of my curriculum vitae which is attached as an Exhibit 1003.

III. MATERIALS CONSIDERED AND SUMMARY OF OPINIONS

10. The opinions contained in this Declaration are based on the documents I reviewed, my professional judgment, as well as my education, experience, and/or knowledge regarding technologies relating to, among other things, video or image signal processing.

11. In forming my opinions expressed in this Declaration, I reviewed the '005 patent (Ex. 1001); the prosecution file history for the '005 patent (Ex. 1004); Hanami et al., U.S. Patent No. 6,122,317 (Ex. 1005); Mitchell et al., MPEG Video Compression Standard, NY: Chapman & Hall, 1996 (Ex. 1006); U.S. Patent No. 5,745,183 to Lam (Ex. 1007); Certified English-language translation of Japanese Patent Application Publication No. H7-250328 to Ishihara et al., Original Japanese Patent Application Publication No. H7-250328, and Translation Certificate (Ex. 1008); Ishihara et al., "A Half-pel Precision MPEG2 Motion-Estimation Processor with Concurrent Three-Vector Search," *IEEE Journal of Solid-State Circuits*, Vol. 30, No. 12, pp. 1502-1509 (Dec. 1995) (Ex. 1009); Information Technology – Generic Coding for Moving Picture and Associated Audio – Recommendation H.262 (ISO/IEC 13818-2), International Organization for Standardisation (ISO/IEC JTC1/SC29/WG11) (March 25, 1994) (Ex. 1010); Wiegand *et al.*, "Special Section on the Joint Call for Proposals on High Efficiency Video Coding (HEVC)

Standardization,” *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 20, No. 12, pp. 1661-1666 (December 2010) (Ex. 1011); U.S. Patent No. 5,905,542 to Linzer (Ex. 1012); and any other materials I refer to in this Declaration in support of my opinions.

12. All of the opinions contained in this declaration are based on the documents I reviewed and my knowledge and professional judgment. My opinions have also been guided by my appreciation of how a person of ordinary skill in the art (as defined below in Section IV) would have understood the claims and the specification of the '005 patent at the time of the alleged invention. My opinions reflect how one of ordinary skill in the art would have understood the '005 patent, the prior art to the patent, and the state of the art at the time of the alleged invention.

13. Based on my experience and expertise, it is my opinion that certain references disclose all the features recited in claims 1-3, 5, 6, 8, 39 and 41 (“challenged claims”) of the '005 patent, as I discuss in detail below.

IV. PERSON OF ORDINARY SKILL IN THE ART

14. I have been asked to consider the time of the alleged invention for the '005 patent to be the late 1990s time frame, including April 30, 1999, which is the filing date of the patent application corresponding to the '005 patent. (Ex. 1001 at Cover.)

15. I am familiar with the level of ordinary skill in the art with respect to the video coding technologies disclosed and claimed in the '005 patent in that timeframe. Based on my review of the '005 patent, the technology, the educational level and experience of active workers in the field of video coding technologies, the types of problems faced by workers in the field, the solutions found to those problems, the sophistication of the technology in the field, and drawing on my own experience, I believe a person of ordinary skill in the art at that time would have had at least a Bachelor's degree in Electrical Engineering, Computer Science, or the equivalent thereof, and two or more years of experience with data compression systems and algorithms, including video coding. More education can supplement practical experience and vice versa. Significantly more practical experience could also qualify one not having the aforementioned education as a person of ordinary skill in the art while, conversely, a higher level of education could offset a lesser amount of experience.

16. I provide my analysis of the '005 patent, the prior art, and my opinions in this declaration from the perspective of a person of ordinary skill in the art, as I have defined it above, during the relevant time frame stated above, where appropriate.

V. TECHNICAL BACKGROUND

17. In this section, I discuss the state of the art relevant to the subject matter of the '005 patent as would have been known, understood, and appreciated by a person of ordinary skill in the art at the time of the alleged invention for the '005 patent.⁵ The '005 patent relates “generally to digital video compression, and, more particularly, to a motion estimation method and search engine for a digital video encoder.” (*See* Ex. 1001, 1:6-8.) Further, the motion estimation method “permits concurrent motion estimation using multiple prediction modes.” (*Id.* at 1:10-11).

18. While my discussion below relates generally to the concepts disclosed and discussed in the '005 patent, I particularly focus on the concepts related to challenged claims 1, 2, 3, 5, 6, 8, 39 and 41 as they were known and understood at the time of the alleged invention. As I discuss below and in section IX, all of the elements set forth in these claims were well-known before the alleged invention of the '005 patent.

19. To begin, I discuss the concept of moving pictures (video) and show why video compression is important when dealing with video data. Then I introduce

⁵ The references I cite in this section support my understanding of what was widely known at the time of the alleged invention.

some key concepts associated with video compression including: (1) reducing spatial and temporal redundancy and (2) encoders and decoders in some typical video encoding systems, including the general process of taking raw image data and compressing and decompressing it. Later, I summarize some of the technical concepts for a conceptual understanding of the '005 patent, including pixel arrays, motion estimation, prediction modes, and concurrent (or parallel) motion estimation methods using multiple prediction modes.

A. Overview of Video Compression/Video Encoding

20. In this section, I provide a high-level overview to explain some concepts associated with video compression. While the details differ from standard to standard, the concepts I explain here apply to many of the standardized hybrid video coding systems (e.g., MPEG-1, MPEG-2) known at the time of the alleged invention.⁶ As described by the *Mitchell* textbook entitled “MPEG Video Compression Standard” (Ex. 1006), which was (and continues to be) an authoritative and popular textbook regarding MPEG at the time of the alleged invention (see below Section VIII.B), a video is composed of many sequential pictures along with

⁶ For several of the concepts I discuss below in this section, I also provide how these concepts are disclosed by *Hanami* in section VIII.A below.

corresponding audio for every picture. For example, “[a] sequence of pictures with accompanying sound track can occupy a vast amount, of storage space when represented in digital form.” (Ex. 1006 at 37 (Section 1.1).)⁷

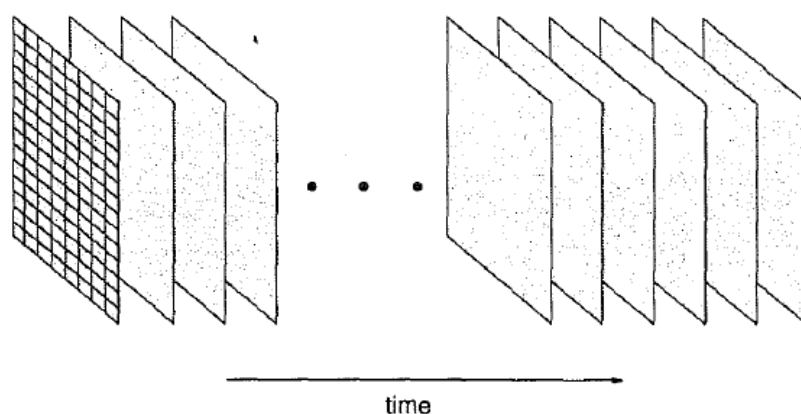


Figure 1.1: Illustration of an MPEG video sequence.

(*Id.* at 40, Figure 1.1: Illustration of an MPEG video sequence.)

21. *Mitchell* exemplifies the size of an uncompressed video, comprising of sequence of pictures and accompanying audio:

For example, suppose the pictures in a sequence are digitized as discrete grids or arrays with 360 pels (picture elements) per raster line and 288 lines/picture, a resolution

⁷ For Exhibit 1006, I cite to the exhibit page numbers appended to the exhibit.

that is fairly typical for MPEG-1. Assuming the picture sequence is in color, a three-color separation (equivalent to the three primaries, red, green, and blue) can be used for each picture. If each color component in the separation is sampled at a 360x288 resolution with 8-bit precision, each picture occupies approximately 311 Kbytes. ***If the moving pictures are sent uncompressed at 24 pictures/second, the raw data rate for the sequence is about 60 Mbit/s, and a one-minute video clip occupies 448 Mbytes.***

(*Id.* at 37 (Section 1.1) (emphasis added).)

22. Video compression can be performed on an uncompressed video to reduce the total size of the video and/or convert the video in a suitable format that can be stored and later accessed. For example, *Mitchell* describes the issue that video compression “addresses is how to store a moving picture sequence, including the audio information, in a compressed format suitable for a digital storage medium such as CD-ROM (compact disc - read only memory).” (*Id.*) Because video compression allows storing video by occupying small memory size, it found application in different industries. And “[w]ith the sharing of technology among [] diverse industrial interests, standardization was a natural development.” (*Id.* at 38 (Section 1.2).) It was desired that “[i]f standardization of digital video could be achieved, the potentially large sales volumes in the consumer market would lower the cost of VLSI

chips for all of the industries. Thus, the cost reduction implicit in having a common digital video technology for all has been a key driving force for standardization.” (*Id.*) Thus video compression (or video encoding) methods were standardized to find application across different industries.

23. Before the alleged invention of the '005 patent, video compression standards, such as MPEG-1 and MPEG-2, were well-known and widely available. (*Id.* at 9 (Foreword); *see below* Section VIII.B.) Each of these standards generally included receiving an input signal (e.g., a raw data signal) representing a sequence of pictures or frames, partitioning each frame into a set of contiguous and nonoverlapping blocks (e.g., 16 x 16), performing motion estimation (ME) on each of these blocks relative to one or more reference frames as specified, finding the difference between the block being encoded and its ME match from the reference frame, further partitioning that difference into smaller blocks (e.g., 8 x 8) applying an orthogonal transformation to the block (e.g., DCT), quantizing the transformed signal, and applying variable length coding (e.g., run-length and Huffman coding) to the quantized values. As mentioned, this process of video encoding is performed on a block-by-block basis where each picture or frame is divided into a plurality of blocks.

24. Motion estimation (ME), which is addressed by the '005 patent, matches blocks of pixels in the current video picture (or frame) being encoded with blocks in a reference picture (e.g., a preceding frame) so that if a suitable match is found, instead of coding and transmitting the contents of the block, a motion vector in addition to the content of the difference between the current block and its match can be encoded and transmitted with a reduced number of bits. (*See, e.g.,* Ex. 1001 at 2:17-3:6.) The motion vector represents the matching block's position (displacement) in the reference frame relative to the block in the current frame.

B. Pixel Arrays and Macroblocks

25. A pixel is the smallest component of a picture, and is also called a "picture element" or pel. (Ex. 1006 at 37 (Section 1.1), 41 (Section 1.3.3).) "[A] picture is made up of a two-dimensional (2-D) grid" arrangement of pixels. (*Id.* at 39 (Section 1.3.1).) ***The macroblock is a basic building block of an MPEG picture*** (shown below in Figure 2.6) and "consists of a 16 x 16 sample array of luminance (grayscale) samples together with one 8 x 8 block of samples for each of two chrominance (color) components. The 16 x 16 sample array of luminance samples is actually composed of four 8 x 8 blocks of samples, and these 8x8 blocks are the units of data that are fed to the compression models." (*Id.* at 58 (Section 2.3.3).)

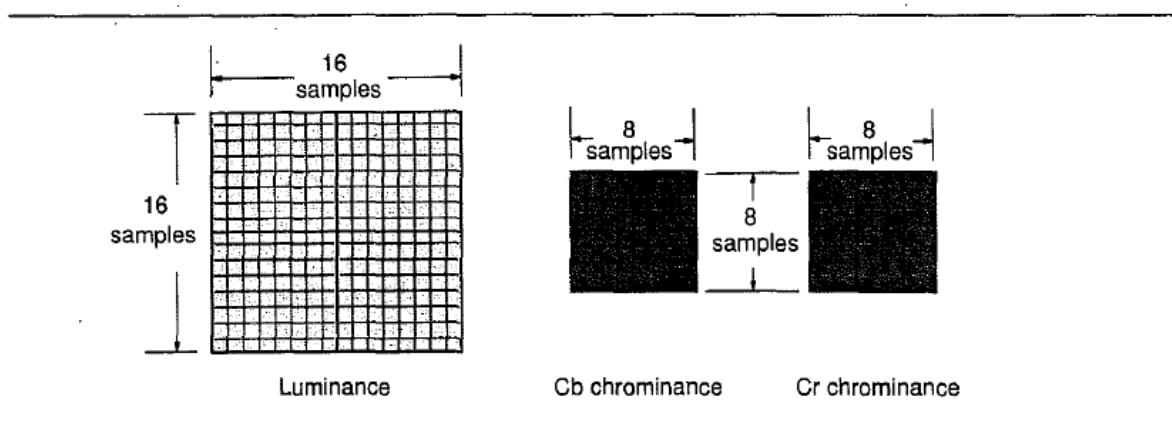


Figure 2.6: The MPEG macroblock.

(*Id.* at 59, Figure 2.6: The MPEG macroblock.)

C. Motion Estimation

26. “The pictures in moving picture sequences often have large regions in which very small changes occur from one picture to the next in the sequence.” (*Id.* at 269 (Section 11.1).). This kind of motion is observed in, for example, videos where an object changes direction with minimum change in the background of the object. In such instances, “[t]he picture-to-picture correlation in these regions is very high and any reasonable video compression scheme should take advantage of this. Obviously, there are usually some changes from picture to picture, or we would only have to send the first picture of the sequence. Changes are caused by a number of different mechanisms, among them noise, scene lighting, and motion.” (*Id.*)

27. “If there is motion in the sequence, a better prediction is often obtained by coding differences relative to areas that are shifted with respect to the area being coded, a process known as motion compensation. *The process of determining the motion vectors in the encoder is called motion estimation.*” (*Id.* at 63 (Section 2.3.7) (emphasis added).) Motion vectors are helpful in determining the change in the position of an object with respect to its previous position. “The motion vectors describing the direction and amount of motion of the macroblocks are transmitted to the decoder as part of the bitstream.” (*Id.* at 63-64 (Section 2.3.7).)

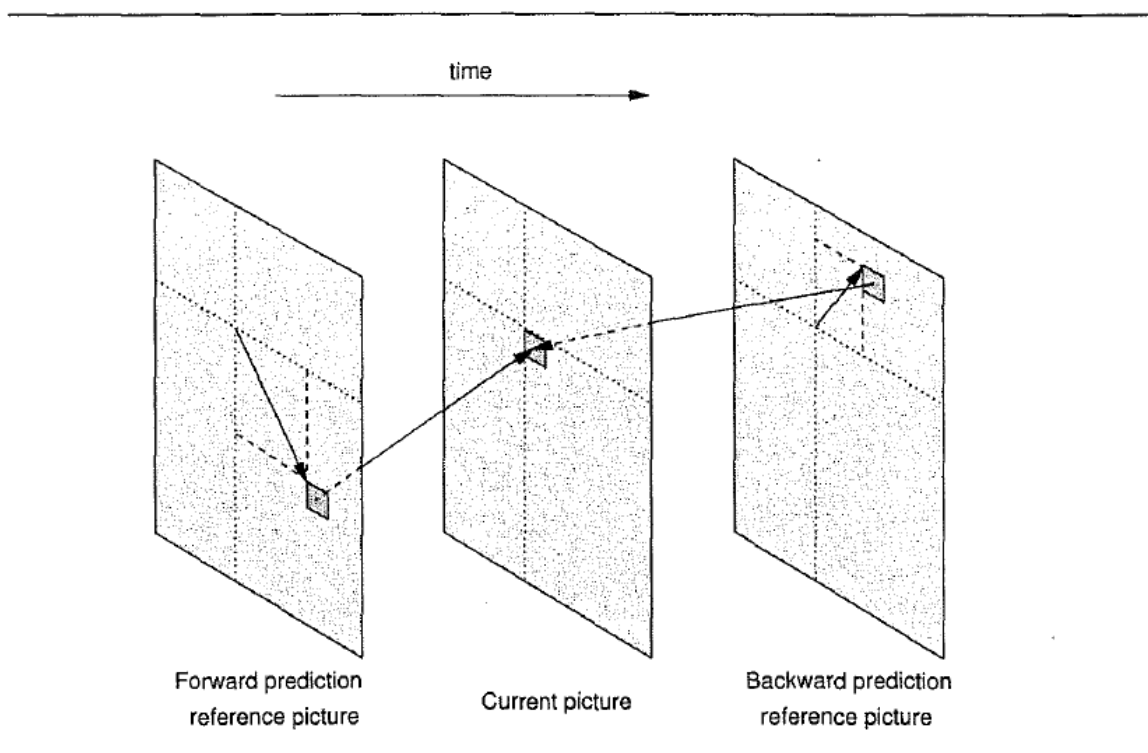


Figure 2.9: Motion compensated prediction and reconstruction.

(*Id.* at 63, Figure 2.9: Motion compensated prediction and reconstruction.)

28. As shown above in Figure 2.9, “[t]he motion vector and corresponding vertical and horizontal displacements are shown for forward and backward motion compensation. The vectors are the same for every pel in the macroblock, and the vector precision is either to full pel or half-pel accuracy.” (*Id.* at 64.) A *pel* is a picture element. (*Id.* at 41 (Section 1.3.3).)

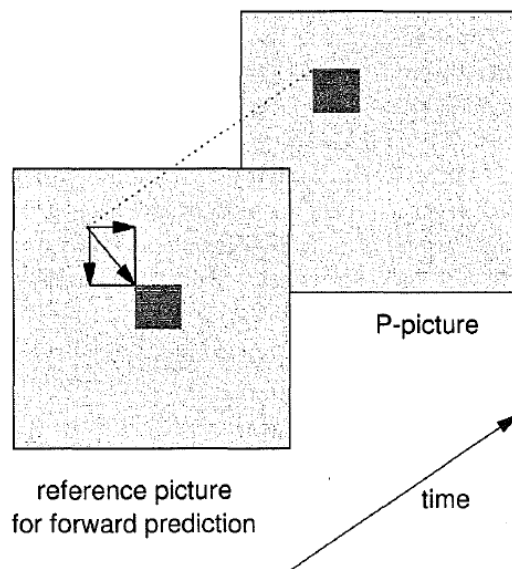


Figure 11.1: P-picture motion vector displacements. Positive displacements in the reference picture are to the right and down, relative to the macroblock being coded.

(*Id.* at 271, Figure 11.1: P-picture motion vector displacements. Positive displacements in the reference picture are to the right and down, relative to the macroblock being coded.)

29. As shown in the Figure 11.1 above, “[i]n each still picture the motion can be described by a two-dimensional array of motion vectors that give displacements relative to a reference picture in the sequence. If good values of these displacements are available, pels in the reference picture can be used to predict the current picture.” (*Id.* at 270 (Section 11.1).)

30. Another example shows B-picture motion vector displacements that has both forward and/or backward motion-compensated prediction.

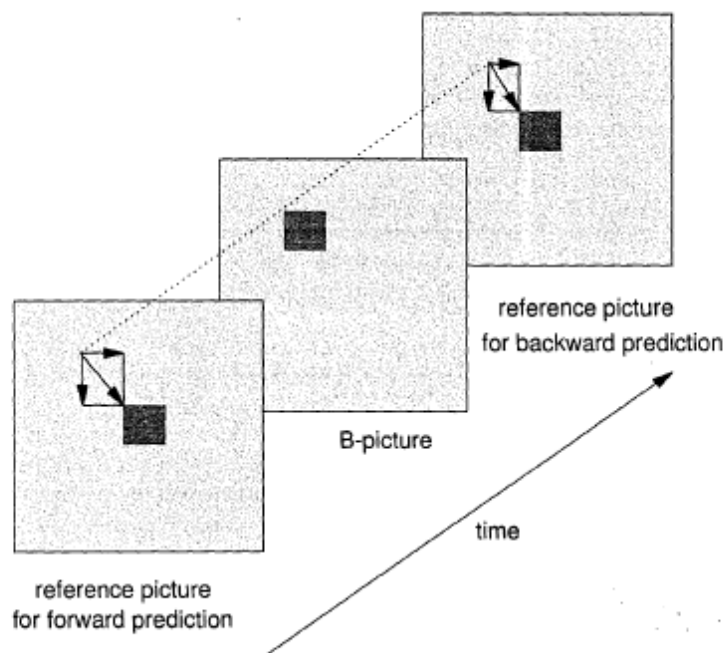


Figure 11.2: B-picture motion vector displacements. In both reference pictures positive displacements are to the right and down, relative to the macroblock being coded.

(*Id.* at 271, Figure 11.2: B-picture motion vector displacements. In both reference pictures positive displacements are to the right and down, relative to the macroblock being coded.)

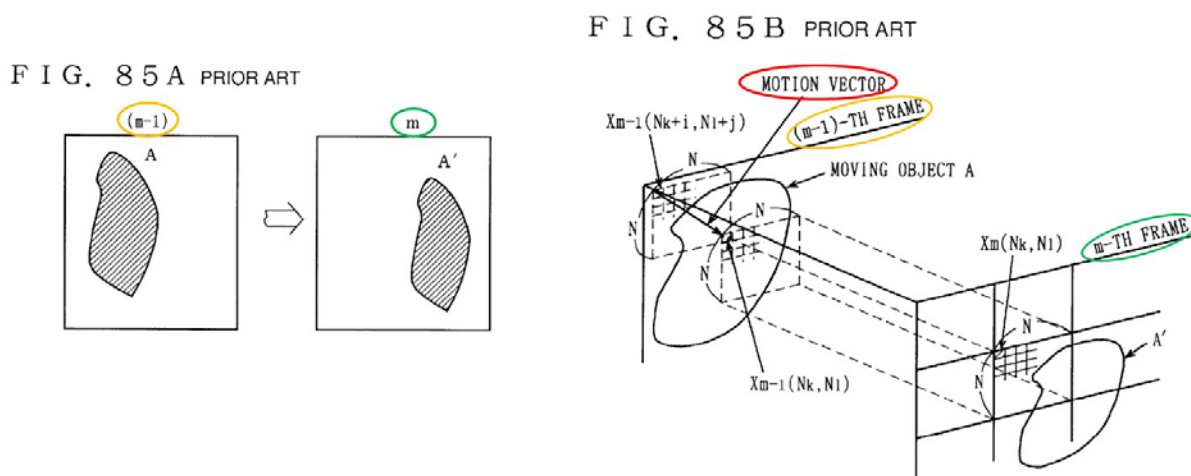
31. *Mitchell* describes that “[m]otion estimation is absolutely necessary in any complete implementation of an MPEG encoder, but the procedures for motion estimation are not normative.” (*Id.* at 314.)

32. With respect to the difference between motion compensation and motion estimation, *Mitchell* discloses that “[m]otion compensation refers to the use of motion displacements in the coding and decoding of the sequence; motion estimation refers to the determination of the motion displacements.” (*Id.*) “The real role of motion compensation is to improve the coding efficiency by improving the prediction from the reference picture. Prediction is always followed by further information, even if this information only serves to tell the decoder that the prediction requires no corrections.” (*Id.* at 270 (Section 11.1).)

D. Block Matching

33. Block matching is a method of reducing temporal correlation by locating regions of video frames or fields that closely match each other on a block-by-block basis during motion estimation. The horizontal and vertical displacements between each pair of matched blocks forms the components of a motion vector.

34. For example, as I discuss below in Section VIII.A.3, *Hanami* discloses the “[c]omput[ation] of a motion vector [] as to interframe predictive coding,” where “[i]n general, **block matching** is employed for motion vector computing.” (Ex. 1005 at 2:65-67 (emphasis added).)



(*Id.* at FIGS. 85A-85B (annotated).)

35. As shown in the figures above, Figures 85A-85B demonstrate the encoding of a block in a frame *m* (annotated in green) where an object moves from position A in frame *m-1* (annotated in orange) to a position A' in frame *m*. (*Id.* at 3:1-2; 3:10-11.) *Hanami* explains the concept of “evaluation values” and “motion vectors” with reference to these Figures, as described below:

In **block matching**, the screen (one frame in this case) is split into blocks of P by Q pixels (P=Q in general). A block most approximate to that of interest in the current frame is found out from a preceding frame . . .

. . . The absolute value of the difference between the data of the corresponding pixels in a block of the preceding frame ((m-1)-th frame) whose pixel position is displaced by a vector (i, j) and in the block in the current frame (m-th frame) is obtained. Then, the displacement vector (i, j) indicating the displacement is changed to various values, and absolute differential sum is obtained for the respective values. *The absolute differential sums are generally called evaluation values. A displacement vector (i, j) indicating the position providing the minimum absolute differential sum is defined as a motion vector.*

(*Id.* at 3:2-25 (emphasis added).)

36. That is, in the process of block matching for motion vector coding, the differential signals between the current block being coded and the candidate blocks in a reference frame are obtained. These differential signals are represented by absolute differential sums—these absolute differential sums are referred to as “*evaluation values*.” The displacement vector associated with the relative position of the current block and the reference block with the minimum evaluation value is defined as a “*motion vector*” (annotated in red above) for the current block being coded.

37. The '005 patent also confirms this well-known understanding of block-matching as it discloses that

For all MPEG-2 prediction modes, *the fundamental technique of motion estimation consists of comparing the current macroblock with a given 16-by-16 pixel array in the anchor picture, estimating the quality of the match according to the specified metric, and repeating this procedure for every such 16-by-16 pixel array located within the search range.* The hardware or software apparatus that performs this search is usually termed the “search engine,” and there exists a number of well-known criteria for determining the quality of the match. Among the best-known criteria are the Minimum Absolute Error (MAE), in which the metric consists of the sum of the absolute values of the differences of each of the 256 pixels in the macroblock with the corresponding pixel in the matching anchor picture macroblock; and the Minimum Square Error (MSE), in which the metric consists of the sum of the squares of the above pixel differences. In either case, *the match having the smallest value of the corresponding sum is selected as the best match within the specified search range, and its horizontal and vertical positions relative to the current macroblock therefore constitute the motion vector.* If the resulting minimum sum is nevertheless deemed to[o] large, a suitable match does not exist for the current macroblock, and it is coded as an intra macroblock. For the purposes of the present invention, either of the above two criteria, or any other suitable criterion, may be used.

(Ex. 1001 at 2:49-3:6.)

E. Prediction Modes and Selection of an Optimum Prediction Mode

38. The MPEG-2 Part 2 standard describes six different prediction modes for motion estimation, three of which apply to frame pictures and three of which

apply to field pictures. (Ex. 1010 (MPEG-2 Part 2 standard) at 82-84 (Section 7.6.1 entitled “Prediction modes”)⁸; *see also id.* generally at 81-95 (Section 7.6 entitled “Motion compensation”).)⁹ An article by Ishihara *et al.* entitled “A Half-pel Precision MPEG2 Motion-Estimation Processor with Concurrent Three-Vector Search” (Ex. 1009) published in a 1995 IEEE Journal summarizes these prediction modes in Table 1:

⁸ For Exhibit 1010, I cite to the exhibit page numbers appended to the exhibit.

⁹ I also discuss the MPEG-2 standard as shown in Exhibit 1010 and its publication date of 1996 in Section IX.A.4 below.

TABLE I
PREDICTION MODES IN THE MPEG2

Picture structure	Prediction mode	Template block size (number of vector)
Field	Field-based prediction	16x16 (one vector)
	16 x 8 MC	16x8 upper, 16x8 lower (two vectors)
	Dual-prime	16x16 (one vector and dmV)
Frame	Frame-based prediction	16x16 (one vector)
	Field-based prediction	16x8 top, 16x8 bottom (two vectors)
	Dual-prime	16x16 (one vector and dmV)

(Ex. 1009 at 6 (Table 1).)¹⁰

39. As I discuss below in Section VIII.A.4, *Hanami* also discloses in its background section the well-known concept of performing motion estimation in a plurality of different predictive modes in order to determine an optimum prediction mode. (Ex. 1005 at 6:1-37.) *Hanami* specifically discloses that “[i]n order to attain excellent coding efficiency (to reduce the quantity of coded data), it is necessary to

¹⁰ For Exhibit 1009, I cite to the exhibit page numbers appended to the exhibit.

perform motion detection in accordance with a plurality of predictive image detection systems (predictive coding systems), thereafter select the optimum predictive image detection system, and detect motion vectors in accordance with the optimum predictive image detection system.” (*Id.* at 6:3-10.) *Hanami* goes on to disclose the details of the predictive modes (predictive coding systems) as follows:

Fields or a frame can be employed as a unit forming a screen. A single frame is formed by two fields (even and odd fields). For example, there are the following predictive image detection systems for the respective cases:

(A) In case of coding pixel data in units of fields:

(a) A field image is split into a plurality of blocks each of P by Q pixels, and a single motion vector is detected (a single predictive image is generated) for each block.

(b) Each block is further split into two blocks along the vertical direction on the screen, and a single motion vector is detected for each of the two blocks. Therefore, motion vectors for the upper and lower blocks are detected (two predictive images are generated) for each block of P by Q pixels.

(B) In case of coding pixel data in units of frames:

(a) A frame image is split into a plurality of blocks each of P by Q pixels, and a single motion vector is detected (a single predictive image is generated) for each block.

(b) Each block of P by Q pixels is split into two pixel groups of those belonging to the same fields, i.e., even and odd fields respectively, and a single motion

vector is detected for each pixel group. Therefore, a motion vector for the even field pixel group and that for the odd field pixel group are detected for each block of P by Q pixels (two predictive images are generated for a single block).

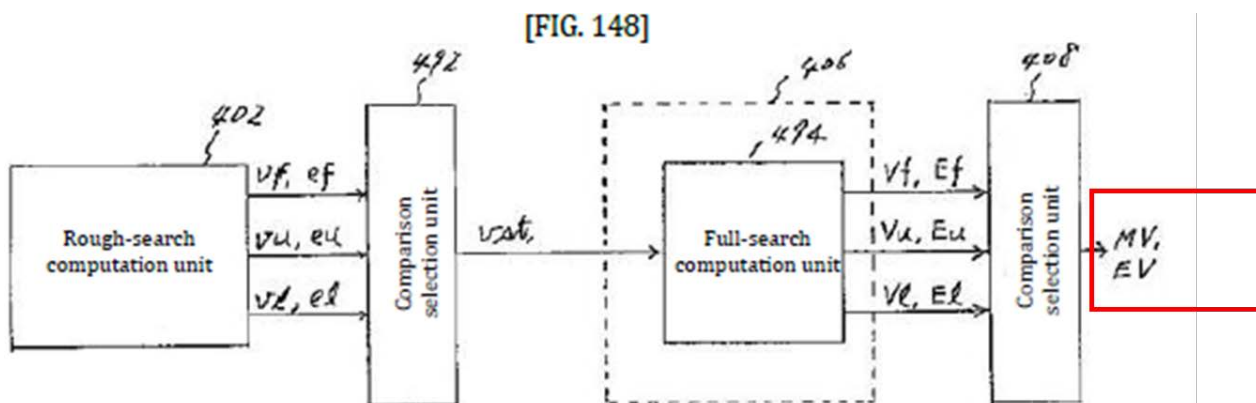
As a structure for executing *the plurality of predictive coding systems (predictive image detection systems)*, it is *desirable to efficiently code the detected motion vectors in any predictive image detection system, in order to attain an effect of improving the coding efficiency.*

(Ex. 1005 at 6:10-46 (emphasis added).)

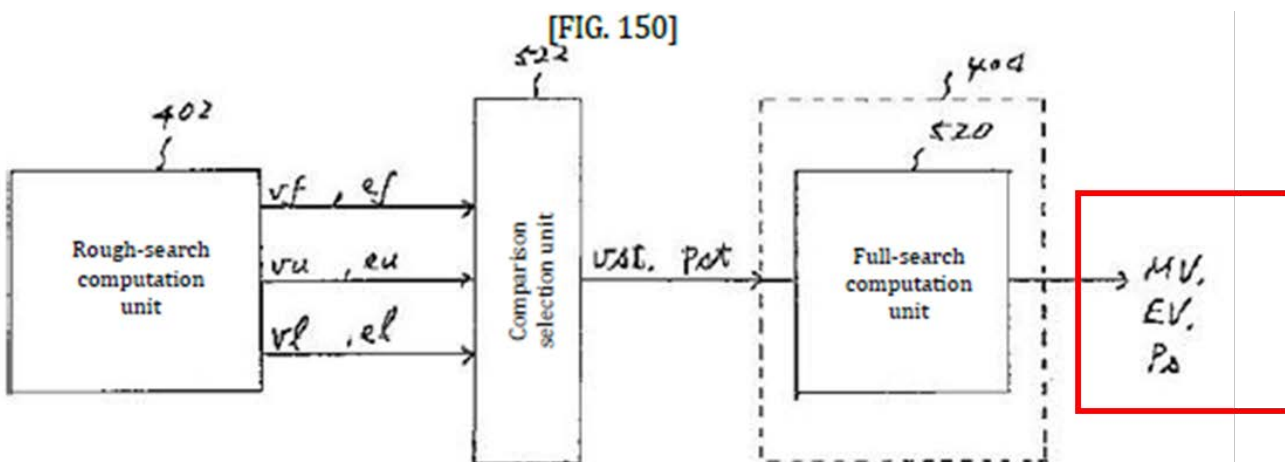
40. As noted above, the option to perform motion estimation according to a plurality of prediction modes was well-known prior to the alleged priority date of the '005 Patent and was included in the specification of the MPEG-2 standard (Ex. 1010). It was further well-known that the primary purpose of being able to perform motion estimation for each block according to a plurality of prediction modes in MPEG-2 is to be able to select the prediction mode that results in the best match among all the possible prediction modes for the block being encoded, that is, to perform motion estimation according to an optimum prediction mode.

41. For example, a 1995 Japanese patent publication ("*Ishihara JP '328*") (Ex. 1008) discloses that "[t]he *comparison selection unit 408 determines an optimal prediction mode* from the evaluation function values Ef, Eu, and El imparted from the full search computation unit 494, and outputs a displacement vector and

evaluation function value corresponding to this **optimal prediction mode as the motion vector MV** and the evaluation function value EV.” (Ex. 1008 at ¶[0383] (emphasis added); see also *id.* at ¶[0357], [0369], [0374], [0383].)



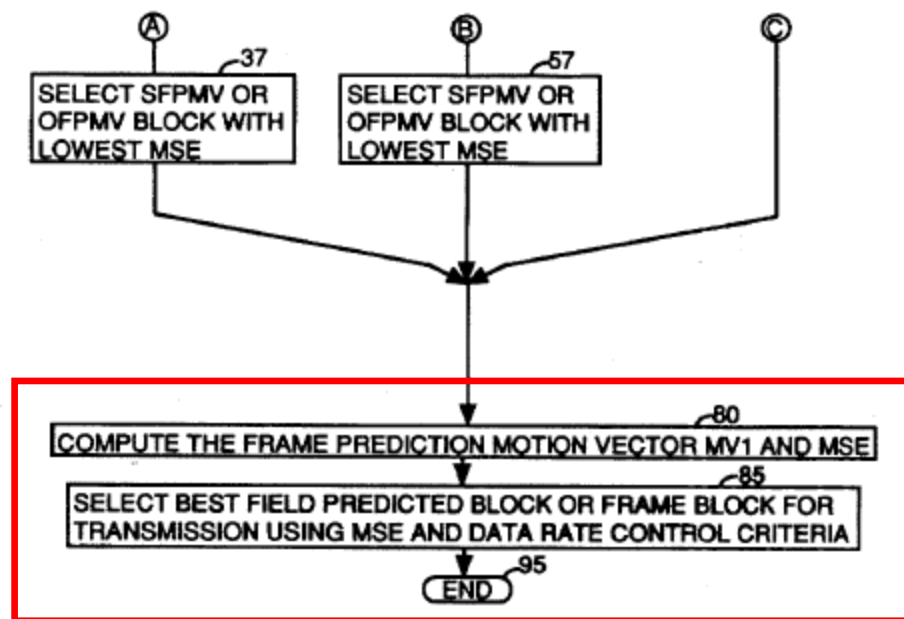
(*Id.* at 229, FIG. 148 (annotated).)



(*Id.* at 230, FIG. 150 (annotated).)

42. In another example, a US patent to *Lam* (Ex. 1007) filed August 25, 1995 and issued in 1998 (Ex. 1007) discloses method for compressing video information selected from a plurality of motion dependent predictions. *Lam* discloses that “[s]everal candidate predicted macroblocks may be formed for each macroblock currently being predicted by using, for example, frame based, field based and ‘dual prime’ encoding. The best prediction for a particular Current Macroblock (CM) is selected from these candidate predicted macroblocks to be compressed and transmitted.” (Ex. 1007 at 2:66-3:12; *see also id.* at 4:13-20, FIG. 6.)

FIGURE 6



(*Id.* at 4, Figure 6 (annotated).)

F. Concurrent/Parallel Performance of Motion Estimation In Multiple Prediction Modes

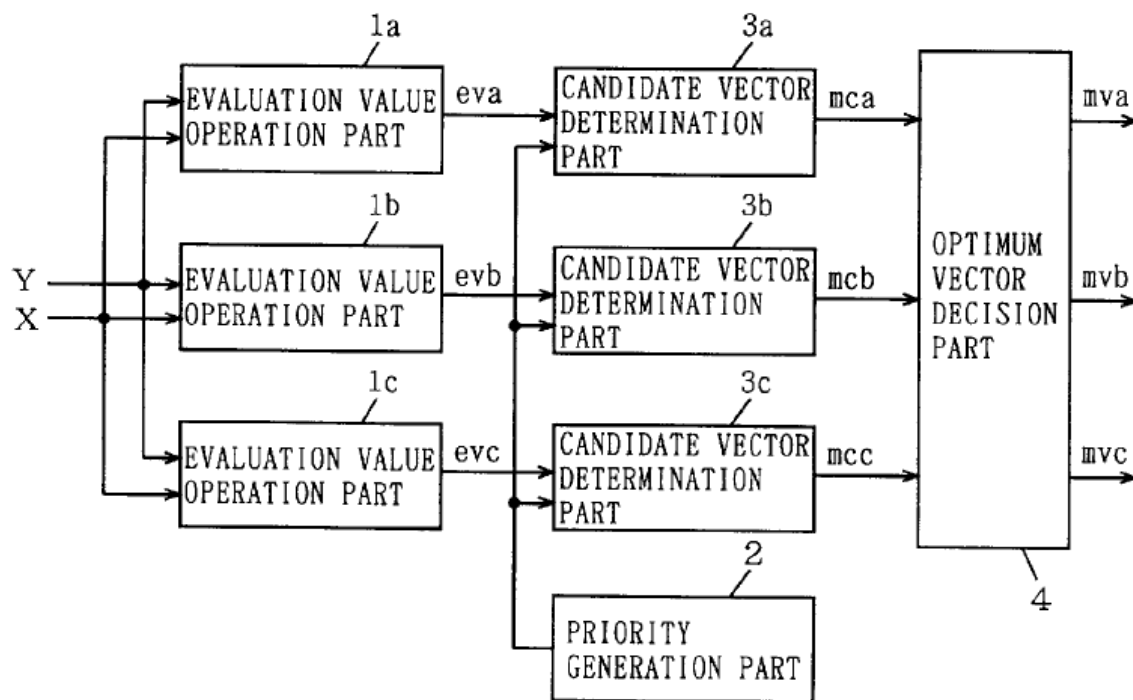
43. The concepts of parallelism and concurrency are ubiquitous in the field of video processing system designs. Parallel computations, operations, architecture have all been used to improve the functionality and efficiency of video processing systems. One such area of video processing systems is the motion estimator engines, which used to perform motion vector searches. In this regard, the concept of concurrently performing motion estimation with respect to a plurality of different prediction modes for a given macroblock being coded was well-known at the time of the alleged invention as discussed below with respect to, e.g., *Hanami* (Ex. 1005), *Ishihara JP'328* (Ex. 1008), and the *Ishihara* IEEE article (Ex. 1009).

44. For instance, as I discuss below in Section VIII.A, just like the '005 patent, *Hanami* discloses an “object of the present invention is to provide a motion vector detector which can *simultaneously detect optimum vectors* efficiently coded for *a plurality of predictive modes in parallel with each other.*” (Ex. 1005 at 6:52-55 (emphasis added).)

45. Even the process by which *Hanami* accomplishes parallel detection of one or more optimum motion vector(s) for a plurality of predictive modes is essentially the same as the '005 patent's technique described above. This is reflected

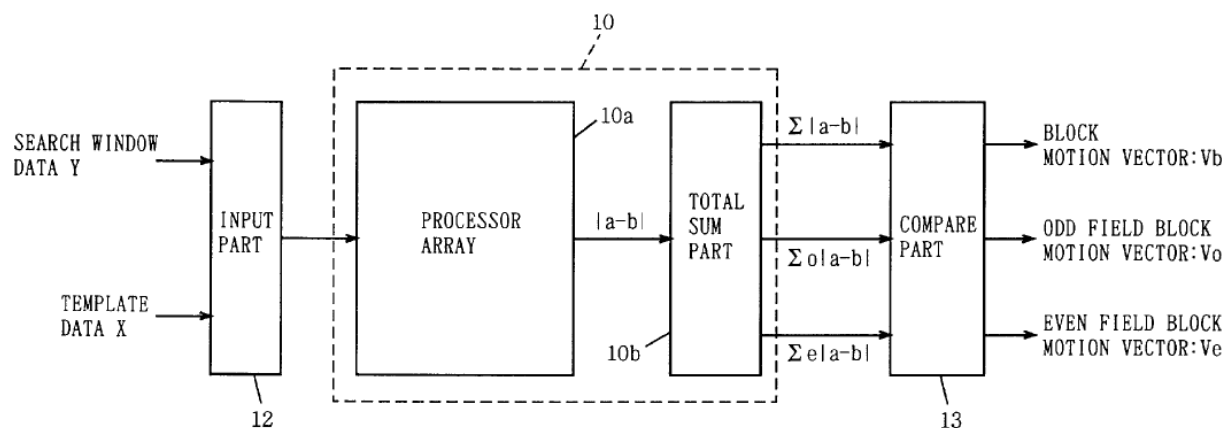
generally by way of Figure 1 and more specifically by way of Figure 4 along with the corresponding description of these figures.

F I G. 1



(*Id.* at FIG. 1.)

F I G. 4



(*Id.* at FIG. 4.)

46. For example, just like the '005 patent, *Hanami* discloses splitting a frame macroblock into two fields (e.g., even and odd fields) and searching each field using field prediction mode against the even and odd fields of a reference picture (i.e., an anchor picture), respectively. Then, the sum of the resulting error metrics is equal to the error of the full macroblock if searched in frame prediction mode. (*Id.* at 13:64-14:9; *see also id.* at 13:29-40, FIG.4.) That is, *Hanami* describes here that for computing evaluation value of a frame predictive mode, evaluation values of the even and odd fields are simply added up with each other which is essentially the same as the '005 patent's description that for computing the error of the frame prediction mode, the error metrics of the top and bottom fields are summed up. (Ex. 1001 at 8:14-44, FIG. 3.)

47. *Ishihara JP'328* (Ex. 1008) similarly discloses parallel processing outputs of multiple prediction modes using summing circuits to generate an evaluation value for each of the prediction modes. In particular, it discloses “[a] summing unit 12 respectively classifies the evaluation-function-value components imparted from the element processors according to a plurality of prediction modes, sums these according to each classification, and generates evaluation-function values for each prediction mode.” (Ex. 1008 at Abstract.) The summing circuits in a computation unit operate in parallel to generate the corresponding evaluation function values for multiple prediction modes simultaneously, “[t]he optimal evaluation function value indicating the most favorable correlation and the corresponding displacement vector are sought from the calculated evaluation function value. The detection of this *optimal evaluation function value and the corresponding displacement vector is executed in parallel in a plurality of respective prediction modes.*” (*Id.* at ¶[0346] (emphasis added); see also *id.* at ¶¶[0350]-[0351], [0359], [0365], [0368], [0371], [0379], FIGS. 138, 144, 148, 150.) “A *comparison unit* 3 determines, according to the evaluation-function values imparted from the summing unit, displacement vectors corresponding to evaluation-function values imparting the *most favorable similarity as motion vectors for each prediction mode.*” (*Id.* at Abstract (emphasis added).)

48. *Ishihara JP'328* (Ex. 1008) further discloses that a comparison unit determines an optimum prediction mode by comparing the evaluation function values corresponding to three prediction modes and selecting the prediction mode with the optimal evaluation-function value. It is explained that “the optimal evaluation point is sought according to the calculated evaluation function values, and a displacement vector corresponding to this optimal evaluation point is determined as a motion vector. ***This full search operation is executed in parallel in a plurality of respective prediction modes.***” (*Id.* at ¶[0350] (emphasis added).) Once the calculation of the evaluation function values for the search window is completed, “the smallest evaluation-function values for each of the three prediction modes are sought, and the displacement vectors corresponding thereto are determined as the motion vector.” (*Id.* at ¶[0140]; *see also id.* at ¶¶[0105], [0152]-[0153].)

49. Similarly, the *Ishihara* IEEE article (Ex. 1009) is directed to an example of an MPEG-2 encoder and discloses “a half-pel precision MPEG2 motion estimation processor ... [that] supports all prediction modes in MPEG2 including frame, field and dual-prime prediction, and estimates three vectors concurrently.” (Ex. 1009 at 1502.) The title and abstract of the *Ishihara* IEEE article indicate that

it discloses “concurrent motion estimation.” For example, the abstract of *Ishihara* explains that it is directed to:

A half-pel precision MPEG2 motion estimation processor using a 0.5 μm CMOS technology ***supports all prediction modes in MPEG2 including frame, field and dual-prime prediction***, and estimates three vectors ***concurrently***.

(Ex. 1009 at Abstract (emphasis added).)

50. The *Ishihara* IEEE article explains that an optimum prediction mode and motion vectors are selected based on concurrently performing motion estimation in a plurality of modes. For example, *Ishihara* states that “[f]or the frame structure pictures, the ***best prediction mode*** and motion vectors for a macroblock are selected from three prediction modes, ***frame-based prediction, field-based prediction, and dual-prime prediction***.” (Ex. 1009 at 1503 (emphasis added).) And that its “requirements for the MPEG2 motion estimation processor are summarized as follows: 1) A capability to estimate motion vectors ***concurrently for various prediction modes***.” (*Id* (emphasis added).) The *Ishihara* IEEE article further discloses that:

In the frame structure pictures, upper-top and lower-top are summed up to estimate the vector for the top 16 x 8 template block, upper-bottom and lower-bottom are for the bottom 16 x 8 template block. Then all groups are summed up for the 16 x 16 template block. These

operations are *performed simultaneously for concurrent three vector search*. The minimum-value detector *finds the minimum distortions and selects the best vectors* for each three template block. In the filed structure pictures, upper-top and upper-bottom are summed up to estimate the vector for the upper 16 x 8 template block, lower-top, and lower-bottom are for the lower 16 x 8 template block. The following operation is the same as for the frame structure pictures. This scheme increases the processing power twice equivalently ((16 x 16) + 2 x (16 x 8) absolute differences per cycle, 20 GOPS at 40 MHz), and realizes a *concurrent three-vector search* with a very small area penalty.

(Ex. 1009 at 1505.)

51. The *Ishihara* IEEE article further explains that concurrently performing motion estimation in a plurality of prediction modes can be done using a chip with a summation circuit operable to output an error metric (mean absolute difference (MAD), which is mathematically similar to the metric of Minimum Absolute Error (MAE) used in the '005 patent) for field mode and frame mode at the same time based on the same search operation. For example, Ishihara discloses a chip referred to as the ME2. (Ex. 1009 at FIG. 5.)

52. Thus, in my opinion, a person of ordinary skill in the art would have been aware of the use of concurrency and parallelism in connection with video coding generally, and particularly with concurrently performing motion estimation in a plurality of different prediction modes, well before the alleged invention of the '005 patent.

VI. THE '005 PATENT

A. Summary of the '005 Patent

53. The '005 patent is directed to techniques for “motion estimation” in a plurality of “prediction modes” in the context of video coding. The '005 patent purports to describe techniques for performing “*concurrent*” motion estimation in a plurality of prediction modes. (Ex. 1001 at 1:6-11.)

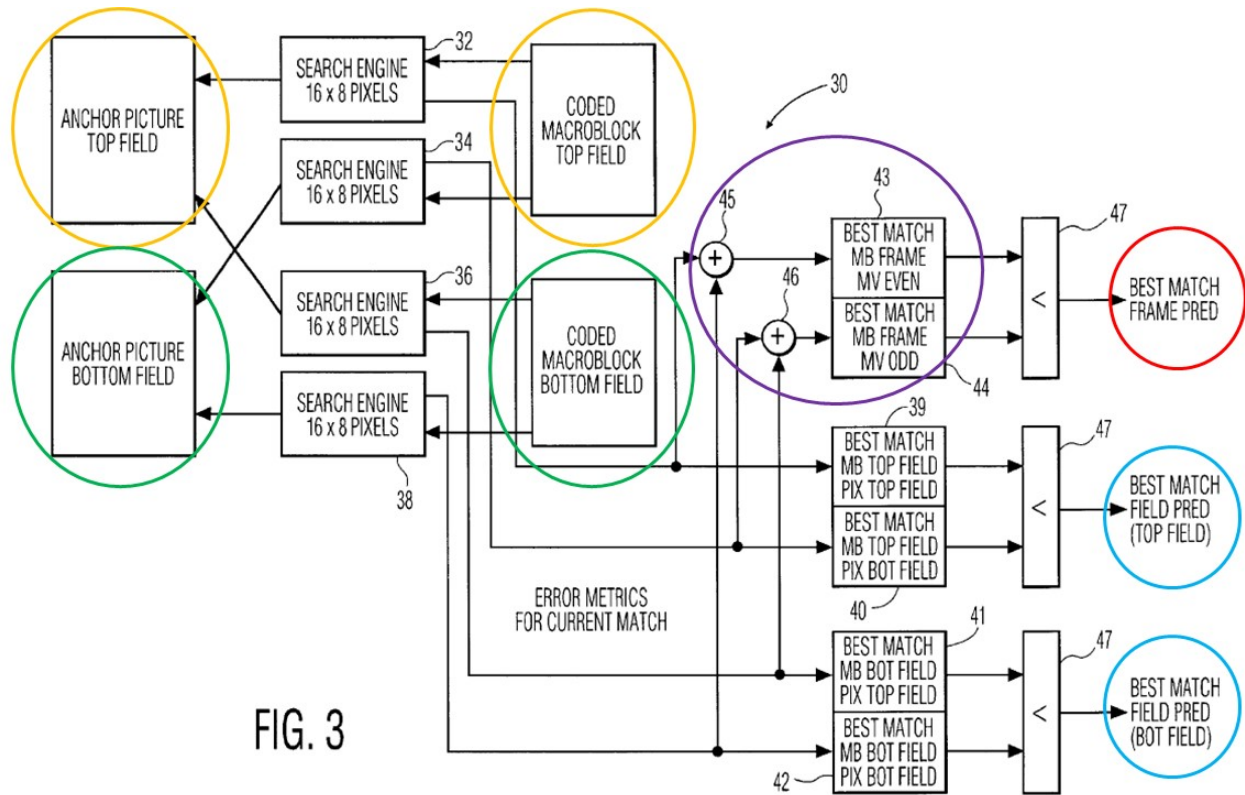
54. The '005 patent concedes that motion estimation was already a well-known video compression technique. (*Id.* at 1:40-55.) The '005 patent similarly concedes that performing motion estimation in a plurality of “prediction modes,” such as the six prediction modes defined by the MPEG-2 standard, was well-known. (*Id.* at 2:49-54, 3:7-13; *see also id.* at 4:24-25 (disclosing that there are “six different possible prediction modes specified by the MPEG-2 standard”).) According to the '005 patent, in conventional methods, a prediction mode “generally” needs to be specified for a given macroblock before motion estimation is performed (the prediction mode is specified based on, e.g., prior knowledge of the picture or image source characteristics). (*Id.* at 3:7-20.) However, it is disclosed that there are many cases where determining the optimum prediction mode would require performing more than one motion estimation for a given macroblock, and doing so is “not a practical option.” (*Id.* at 3:15-24.) The '005 patent purports to solve this problem by

disclosing the allegedly “novel scheme” of performing motion estimation for multiple prediction modes *concurrently*. (*Id.* at 3:54-67.)

55. The specification describes concurrent motion estimation in the context of six prediction modes in the context of frame and field pictures. (*Id.* at 9:36-54, 10:66-11:9.) In particular, as I discuss below, three of these prediction modes are directed to frame pictures and the other three to field pictures. (*Id.* at 4:23-26.) Notably, as I discuss below, the specification describes separate embodiments in which motion estimation is performed in different subsets of the six prediction modes, but it does *not* describe an embodiment in which motion estimation is performed in all six at once.

56. For example, for **frame pictures**, “the motion estimation may be performed, in a frame picture, for both *frame* and *field* prediction modes simultaneously, during the same search of the anchor picture.” (*Id.* at 8:9-13 (emphasis added), FIG. 3.) This is possible because, as shown in an annotated version of Figure 3 below, if the macroblock is split into its constituent fields (*i.e.*, two 16-by-8 arrays) [top field annotated in orange and bottom field annotated in green] and each field is searched using *field prediction mode* against the top and bottom fields of the anchor picture [top field annotated in orange and bottom field annotated in green], respectively, the sum of the resulting error metrics is equal to

the error of the full 16-by-16 block [annotated in purple] if searched in *frame prediction mode*. (*Id.* at 8:14-44.)



(*Id.* at FIG. 3 (annotated); *see also id.* at 8:49-9:35.)

57. This technique allows for the performance of two prediction modes (*frame* prediction mode or *field* prediction mode) for frame pictures based on the same search operation. (*Id.* at 9:27-32.) Of note, while Figure 3 includes 3 “outputs,” these outputs correspond to only two prediction modes: Field prediction mode for frame pictures results in two motion vectors: one for the top field, and one for the bottom field [both annotated in blue above] and frame prediction mode

[annotated in red above]. Dual-prime mode can be examined after performing motion estimation for field prediction, based on the results of the same search operation, by evaluating relationships between the resulting motion vectors.¹¹ (*Id.* at 12:19-42.)

58. For **field pictures**, the '005 patent describes a similar method of concurrently performing motion estimation for (1) field prediction mode for field pictures and (2) 16x8 prediction mode for field pictures. (*Id.* at 9:40-10:65; Fig. 5.) Since macroblocks of field pictures only have a single parity, the macroblock is instead divided into 16x8 blocks comprising “upper” and “lower” halves instead of top and bottom fields as in the case of frame pictures. (*Id.* at 9:40-50.) Dual-prime

¹¹ While the '005 patent specification describes a dual-prime mode, the challenged claims do not require this feature. For example, challenged claims 1, 39, and 41 broadly recite “a plurality of different prediction modes” or “a plurality of different motion estimation prediction modes” while challenged claim 8 sets forth the “different prediction modes” as a “frame prediction mode” and “field prediction mode” for frame pictures.

mode can be performed for field pictures like that for frame pictures.¹² (*Id.* at 12:19-42.)

59. All of the above concepts were well-known before the alleged invention of the '005 patent, as I discussed above in Section V and further discuss below in Section IX.

B. Prosecution History of the '005 Patent

60. I reviewed the prosecution history of the '005 patent and am aware of the following facts. During prosecution, in response to a rejection based on U.S. Patent No. 5,905,542 (“*Linzer*”) (Ex. 1012), the applicant argued that “generating a motion vector in response to the determining step” (claim 1) or “generating one or more motion vectors using the selected prediction mode” (issued claim 41, then claim 50) is not disclosed by *Linzer* because the motion vectors disclosed in *Linzer* are not generated in response to the selection of an optimum prediction mode. (Ex. 1004 at 78-79, 81-90, 94-100.) The Examiner disagreed and noted that *Linzer* discloses producing candidate vectors corresponding to different prediction modes and eliminating candidate vectors based on the prediction mode that is selected, and a final motion vector is generated based on the selected prediction mode (and

¹² See above n.11.

subsequently, this motion vector and the selected prediction mode are output to the downstream encoding circuitry). (*Id.* at 105-107.) The applicant did not dispute the Examiner's position in regard to the "generating" feature(s), and instead newly argued that *Linzer* did not disclose determining an optimum prediction mode "based on which mode is optimum or produces an optimum result," as set forth in some variation in issued claims 1 and 41. (*Id.* at 122.)

61. Without addressing this new argument, the claims were allowed and the reasons for allowance referenced claim limitations that are not recited in challenged independent claims 1, 39, and 41. In particular, the claims were allowed, as the Examiner explained, because they allegedly recite "a method for motion coding a digital video data stream which includes ' . . . determining which of the first, second, third, and fourth error metrics, and first and second composite metrics has the lowest value, and selecting one of the plurality of possible motion estimation prediction modes on the basis of such determination'" (*Id.* at 137.) That is, despite being previously rejected, the challenged claims do not recite these limitations. Instead, these limitations are found in other claims not challenged here.

VII. CLAIM CONSTRUCTION

62. I understand that claim terms are typically given their ordinary and customary meaning, as would have been understood by a person of ordinary skill in the art, at the time of the alleged invention, which I understand is the late 1990s time frame, including April 30, 1999, as I discussed above in Section IV. In considering the meaning of the claims, however, I understand that one must consider the language of the claims, the specification, and the prosecution history of record. I have followed these principles in my analysis of the challenged claims.

63. I have also been asked to consider possible constructions of certain claim terms, as summarized in the table below, and apply these possible constructions in my analysis. Accordingly, I have applied these possible constructions in my analysis and in forming my opinions where appropriate in this Declaration.

Claim Term	First Possible Construction	Second Possible Construction
“concurrently perform[ing]/[s] motion estimation [for]/[using] each of a plurality of different [motion estimation] prediction modes” (Claims 1, 39, 41)	<i>Performing motion estimation according to each of a plurality of prediction modes during the same search operation for a given macroblock, without performing a separate search on the same macroblock for each such mode.</i>	<i>[No second possible interpretation considered]</i>
“optimum” (Claims 1, 39, 41)	<i>Best or most suitable.</i>	<i>At least includes the prediction mode that yields the smallest value of an error metric.</i>

“a best match” (Claim 1)	<i>The second pixel array within a specified search range of the reference picture having the smallest value of an error metric.</i>	<i>At least include the array included in the search having the smallest value of an error metric.</i>
--------------------------	--	--

VIII. OVERVIEW OF THE PRIOR ART¹³

A. *Hanami* (Ex. 1005)

64. *Hanami* discloses techniques for detecting motion vectors employed for motion compensation of moving images (“method for coding an uncompressed digital video data stream”). (Ex. 1005 at 1:5-8; *see also id.* at 1:10-56, 6:48-55, FIGS. 1-4, 83.)

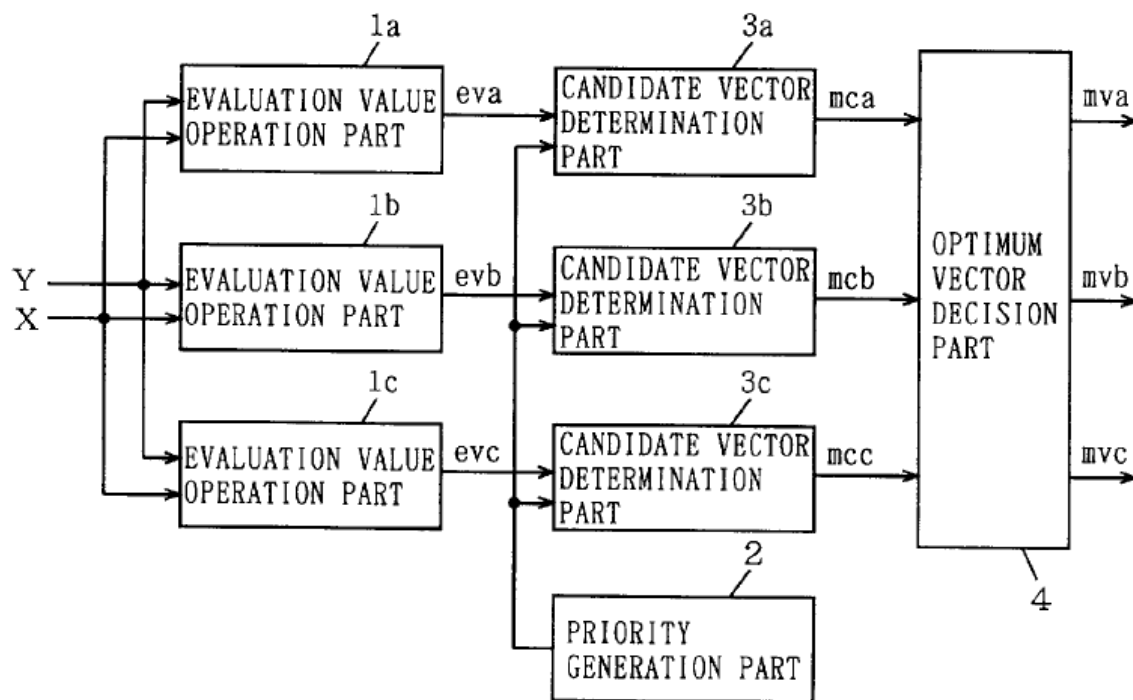
65. In fact, just like the '005 patent, *Hanami* discloses an “object of the present invention is to provide a motion vector detector which can *simultaneously detect optimum vectors* efficiently coded for *a plurality of predictive modes in parallel with each other.*” (*Id.* at 6:52-55 (emphasis added).)

66. Even the process by which *Hanami* accomplishes parallel detection of one or more optimum motion vector(s) for a plurality of predictive modes is basically the same as the '005 patent's technique described above. This is reflected

¹³ These summaries of the prior art in Section VIII are applicable to my analysis below in Section IX. Accordingly, I incorporate these summaries, which reflect how a person of ordinary skill in the art would have understood the references, into my analysis below where appropriate.

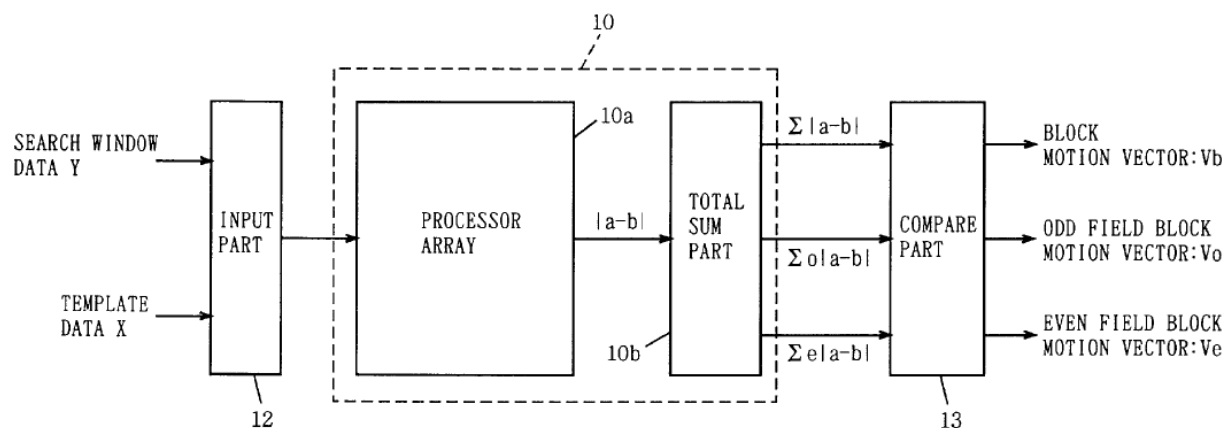
generally by way of Figure 1 and more specifically by way of Figure 4 along with the corresponding description of these figures.

F I G. 1



(Id. at FIG. 1.)

FIG. 4



(*Id.* at FIG. 4.)

67. For example, just like the '005 patent, *Hanami* discloses splitting a frame macroblock into two fields (e.g., even and odd fields) and searching each field using field prediction mode against the even and odd fields of a reference picture (i.e., an anchor picture), respectively. Then, the sum of the resulting error metrics is equal to the error of the full macroblock if searched in frame prediction mode.

A frame includes even and odd fields, and even and odd field pixels are alternately arranged on horizontal scan lines on the screen. Therefore, outputs of the element processors [in processor array 10a of Fig. 4] arranged in correspondence to even and odd field pixels express the components of evaluation values in accordance with the even and odd field predictive modes respectively. ***It is possible to compute evaluation values for the respective ones of the plurality of predictive modes by classifying***

such outputs of the element processors in response to the positions of the element processors arranged in the processor array 10a. In case of the frame predictive mode, the evaluation values of the even and odd fields are simply added up with each other.

(*Id.* at 13:64-14:9 (emphasis added); *see also id.* at 13:29-40, FIG.4.) That is, *Hanami* describes here that for computing evaluation value of a frame predictive mode, evaluation values of the even and odd fields are simply added up with each other which is essentially the same as the '005 patent's description that for computing the error of the frame prediction mode, the error metrics of the top and bottom fields are summed up. (Ex. 1001 at 8:14-44, FIG. 3.)¹⁴

¹⁴ *Hanami*'s even and odd fields correspond to the '005 patent's top and bottom fields, respectively. For example, as noted above in the reproduced portion of *Hanami*, it discloses that "[a] frame includes even and odd fields, and even and odd field pixels are alternately arranged on horizontal scan lines on the screen." (Ex. 1005 at 13:64-66; *see also id.* at 15:54-63, FIGS. 9A-9B (showing even field subtemplate block and odd field subtemplate block).) Similarly, the '005 patent discloses with reference to Figure 1 that "the current macroblock is partitioned into one 16-by-8 array consisting of lines from the top field (even-numbered lines,

68. In fact, a person of ordinary skill in the art would have readily understood from *Hanami*'s figures and description (like the reproduced portions above) that *Hanami*, like the '005 patent, allows for concurrently performing motion estimation in a plurality of prediction modes (e.g., frame prediction mode for frame pictures and field prediction mode for frame pictures) based on the same search operation. (*See also below* Section IX.A.) I provide additional details regarding *Hanami*'s disclosure below.

1. *Hanami*'s Overall Structure and Method

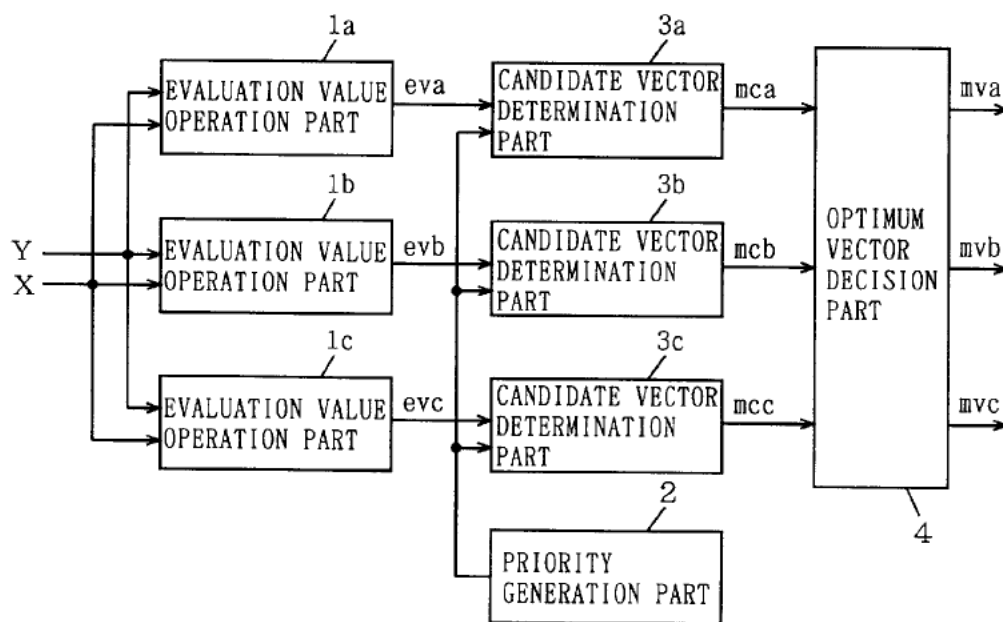
69. As shown below, *Hanami*'s Figure 1 “schematically illustrates the principle structure of a motion vector detector according to the present invention” while Figure 4 “is a block diagram schematically showing the overall structure of a motion vector detector according to an embodiment of the present invention.” (Ex. 1005 at 10:55-56, 12:51-54.) In other words, *Hanami*'s Figure 4 (and its related Figures) disclose a more detailed implementation of the high-level implementation shown in Figure 1. (*Id.*) In fact, *Hanami* explains how certain parts described in Figure 1 relate to Figure 4:

starting with 0), and a second, 16-by-8 array consisting of lines from the bottom field (odd-numbered lines, starting with 1).” (Ex. 1001 at 7:34-38, FIG. 1.)

Referring again to FIG. 1 in relation to the structure shown in FIG. 4, the evaluation value operation parts 1a to 1c correspond to the operation part 10, and the priority generation part 2, the candidate vector determination parts 3a to 3c and the optimum vector decision part 4 correspond to the compare part 13.

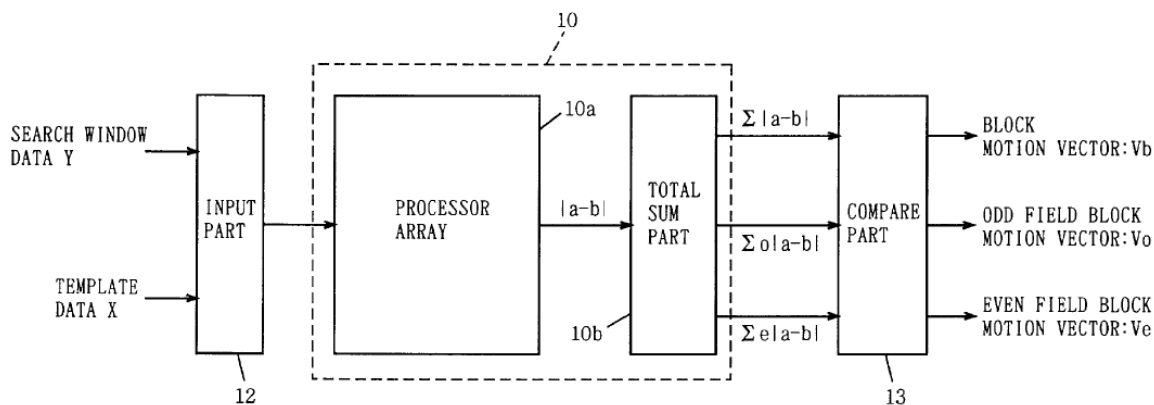
(*Id.* at 13:23-29.)

F I G. 1



(*Id.* at FIG. 1.)

FIG. 4



(*Id.* at FIG. 4.)

Referring to Figure 1, *Hanami* discloses:

the motion vector detector includes evaluation value operation parts 1a, 1b and 1c for receiving template block (current screen image block) pixel data X and search window block (reference screen image block) pixel data Y and [for] computing evaluation values *in accordance with different predictive modes respectively....*

(*Id.* at 10:57-62 (emphasis added); FIG. 1.)

70. *Hanami* discloses that the motion vector detector further includes “an optimum vector decision part 4 for receiving the candidate vectors mca, mcb and mcc from the candidate vector determination parts 3a to 3c in parallel with each other and deciding optimum vectors mva, mvb and mvc serving as motion vectors for the respective predictive modes.” (*Id.* at 11:4-9.) The “[m]otion vector for a

single macro block is decided on the basis of these optimum vectors mva, mvb and mvc.” (*Id.* at 11:9-11.) This is because, as *Hanami* explains, “[i]t is necessary to transmit a motion vector per pixel block.” (*Id.* at 3:26; *see also id.* at 3:39-40 (“Motion vector detection is executed in the units of the blocks.”), 15:38-39 (disclosing that “a motion vector is detected for each macroblock”), 24:3-8.)

71. Similarly, *Hanami* discloses with respect to Figure 4 that “the motion vector detector includes an input part 12 for receiving *search window [] pixel data Y* and *template block pixel data X* included in a prescribed region of a search area and outputting the received data X and Y at prescribed timings respectively.” (*Id.* at 12:56-60 (emphasis added).)

2. Template Block In Current Frame and Search Area In Reference Frame

72. The *template block pixel data X* is, for example, the pixel data of template block 1 (TB1) of a current frame 36 shown in Figures 8A-8B. (*Id.* at 15:38-53; FIGS. 8A-8B; *see also id.* at 14:11-19 (describing an exemplary template block 20 and a search area 22 with reference to FIG. 5); FIG. 5.) The template block TB1 is a block that is currently being encoded.

FIG. 8A

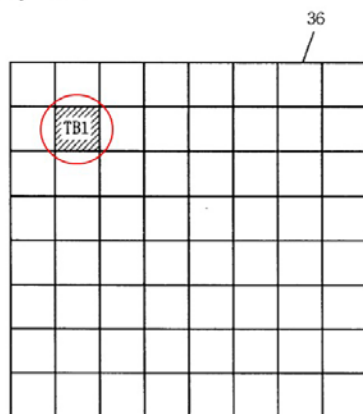


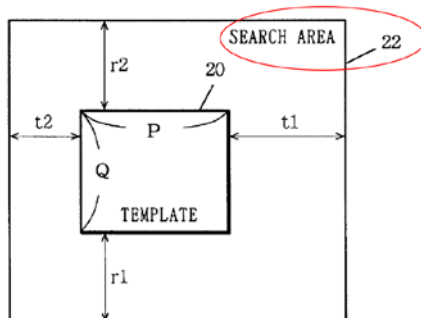
FIG. 8B

a	a		a	a
0,0	0,1		0,14	0,15
a	a		a	a
1,0	1,1		1,14	1,15
⋮				
a	a		a	a
14,0	14,1		14,14	14,15
a	a		a	a
15,0	15,1		15,14	15,15

(*Id.* at FIGS. 8A-8B (annotated).)

73. As shown in Figures 10-11, the *search window pixel data Y* is, for instance, the pixel data representing the collection of macroblocks MB1-MB3 included in a search area of a reference frame 37 (e.g., search area 22 shown in Figure 5), which is stored as data for search window 40. (*Id.* at 14:11-18, 15:64-16:13; FIGS. 5, 10-11.)

FIG. 5



(*Id.* at FIG. 5 (annotated).)

FIG. 10

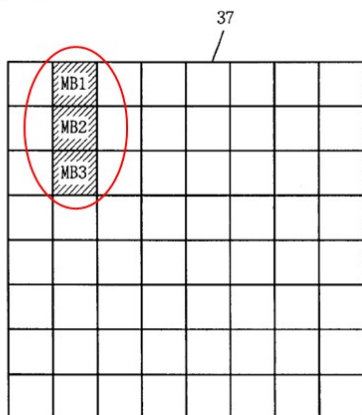
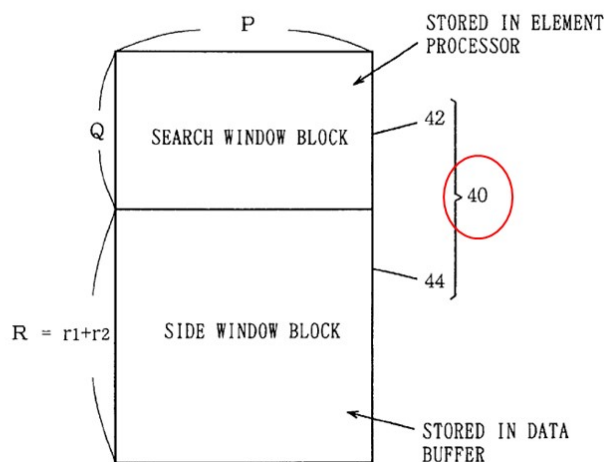


FIG. 11



(*Id.* at FIGS. 10-11 (annotated).)

74. *Hanami* discloses that the search window 40 is a combination of search window block 42 and side window block 44. (*Id.* at 16:14-16; FIG. 11.) And the search window 40 is a subset of the overall search area (e.g., search area 22 in Figure 5) that is used to search for a candidate block with respect to the template block TB1 in the current frame. (*Id.* at 14:11-18, 15:64-16:16, FIGS. 5, 10-11.) Thus, the pixel data included in the search area (for the template block TB1 being coded) includes, but is not limited to, the macroblocks MB1-MB3. (*Id.*) For ease and simplicity of explanation, I refer to the macroblocks MB1-MB3 in my analysis below in the context of the search area/search window 40 used to perform motion estimation with respect to the template block TB1 that is currently being coded.

3. Block-matching, Evaluation Values, and Motion Vector Generation

75. Referring back to Figure 4, the motion vector detector of Figure 4 additionally includes “an operation part 10 for computing three *evaluation values* (e.g., absolute differential sums) $\Sigma|a-b|$, $\Sigma o|a-b|$ and $\Sigma e|a-b|$ as to a displacement vector for a single template block on the basis of the data X and Y supplied from the input part 12.” (*Id.* at 12:60-64 (emphasis added); FIG. 4.) The motion vector detector also includes:

a compare part 13 for receiving the three *evaluation values* $\Sigma|a-b|$, $\Sigma o|a-b|$ and $\Sigma e|a-b|$ obtained in the operation part 10, evaluating the same in accordance with priority assigned to each search window block and deciding *a motion vector* for each predictive mode in accordance with the results of the evaluation.

(*Id.* at 12:65-13:3 (emphasis added).)

76. *Hanami* explains certain basic, fundamental video encoding principles by way of its background section, including the concept of “*evaluation values*” and a “*motion vector*.” (*Id.* at 1:31-55; 2:13-27; 2:65-67; FIGS. 83-85B.) For example, *Hanami* discloses the “[c]omput[ation] of a motion vector [] as to interframe predictive coding,” where “[i]n general, *block matching* is employed for motion vector computing.” (*Id.* at 2:65-67 (emphasis added).)

FIG. 85 A PRIOR ART

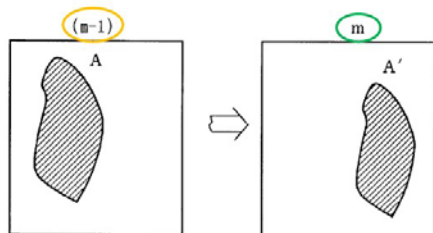
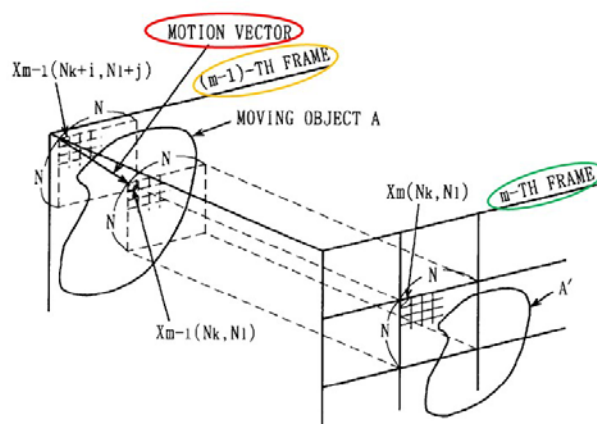


FIG. 85 B PRIOR ART



(*Id.* at FIGS. 85A-85B (annotated).)

77. As shown in the figures above, Figures 85A-85B demonstrate the encoding of a block in a frame m (annotated in green) where an object moves from position A in frame $m-1$ (annotated in orange) to a position A' in frame m . (*Id.* at 3:1-2; 3:10-11.) *Hanami* explains the concept of “evaluation values” and “motion vectors” with reference to these Figures, as described below:

In **block matching**, the screen (one frame in this case) is split into blocks of P by Q pixels ($P=Q$ in general). A block most approximate to that of interest in the current frame is found out from a preceding frame . . .

. . . The absolute value of the difference between the data of the corresponding pixels in a block of the preceding frame (($m-1$)-th frame) whose pixel position is displaced by a vector (i, j) and in the block in the current frame (m -th frame) is obtained. Then, the displacement vector (i, j)

indicating the displacement is changed to various values, and absolute differential sum is obtained for the respective values. *The absolute differential sums are generally called evaluation values. A displacement vector (i, j) indicating the position providing the minimum absolute differential sum is defined as a motion vector.*

(*Id.* at 3:2-25 (emphasis added).)

78. That is, in the process of block matching for motion vector coding, the differential signals between the current block being coded and the candidate blocks in a reference frame are obtained. These differential signals are represented by absolute differential sums—these absolute differential sums are referred to as “*evaluation values*.” The displacement vector associated with the relative position of the current block and the reference block with the minimum evaluation value is defined as a “*motion vector*” (annotated in red above) for the current block being coded.

4. Prediction Modes

79. As I discussed above with reference to Figure 1, *Hanami* discloses that its “motion vector detector includes evaluation value operation parts 1a, 1b and 1c for receiving template block (current screen image block) pixel data X and search window block (reference screen image block) pixel data Y and [for] computing

evaluation values *in accordance with different predictive modes respectively.*” (*Id.* at 10:57-62 (emphasis added); FIG. 1.)

80. In fact, *Hanami* discloses in its background section the well-known concept of performing motion estimation in a plurality of different predictive modes in order to determine an optimum prediction mode. (*Id.* at 6:1-37.) *Hanami* specifically discloses that “[i]n order to attain excellent coding efficiency (to reduce the quantity of coded data), it is necessary to perform motion detection in accordance with *a plurality of predictive image detection systems (predictive coding systems), thereafter select the optimum predictive image detection system, and detect motion vectors in accordance with the optimum predictive image detection system.*” (*Id.* at 6:3-10 (emphasis added).)

81. Like the '005 patent, *Hanami* discloses that the predictive modes can be operated on a frame or a field in a frame. “A single frame is formed by two fields (even and odd fields)” and “[i]n case of coding pixel data in units of frames,” *Hanami* discloses a *frame predictive mode* along with odd/even-*field predictive modes*. (*Id.* at 6:11-13; 6:25-37; 13:4-14; 15:37-63 (emphasis added).)

82. Therefore, as *Hanami* explains with respect to Figures 1 and 4, the three *evaluation values* $\Sigma|a-b|$, $\Sigma_o|a-b|$ and $\Sigma_e|a-b|$ obtained in the operation part 10

represent the absolute differential sums for the frame predictive mode, odd field predictive mode, and even field predictive mode, respectively.

The evaluation value $\Sigma|a-b|$ corresponds to that of a frame predictive mode obtaining the evaluation value using all pixels of a current image block (template block), where a and b represent template block pixel data and corresponding pixel data of a reference image block (search window block) respectively. The evaluation value $\Sigma_o|a-b|$ corresponds to that in accordance with an odd field predictive mode detecting a predictive image using pixels included in an odd field. The evaluation value $\Sigma_e|a-b|$ is that computed in accordance with an even field predictive mode detecting a predictive image using pixels included in an even field.

(*Id.* at 13:4-14.)

83. Subsequently, an optimum motion vector for each predictive mode is obtained by the compare part 13 (e.g., the one representing the minimum evaluation value for each predictive mode represented by V_b , V_o , and V_e in Figure 4). (*Id.* at 12:65-13:3.) Then, a motion vector for the current block is determined from amongst the three obtained optimum motion vectors V_b , V_o , and V_e (e.g., the optimum vector having a minimum evaluation value from amongst the three obtained optimum motion vectors). (*Id.* at 11:9-11, 6:3-10, 15:38-39, 24:3-8.)

B. *Mitchell* (Ex. 1006)

84. *Mitchell* is a textbook titled ‘MPEG Video Compression Standard’ authored by *Joan L. Mitchell, William B. Pennebaker, Chad E. Fogg* and *Didier J. LeGall* and “Digital Multimedia Standards Series” as the series that includes this title. The textbook was published in 1996 by Chapman & Hall of New York. *Mitchell* generally discloses the standardization of video compression by the Moving Picture Experts Group (MPEG) and its various related concepts to help readers “decide what standards to support and use” (Ex. 1006 at 8) and to provide implementers with a “wealth of technical details that help them implement those standards correctly.” (*Id.*)

85. In the foreword section¹⁵, *Mitchell* describes that MPEG was a widely used digital video coding standard because “the merit of MPEG was its ability to identify opportunity and to bring disparate industries to work together.” (*Id.* at 9.) The MPEG standard has been implemented across different industries for example,

¹⁵ The foreword section of *Mitchell* is authored by Leonardo Chiariglione – Convenor, ISO/IEC JT1/SC29/WG11 (MPEG) (*see e.g.*, Ex. 1006 at 8). He is the chairman and co-founded the Moving Picture Experts Group (MPEG) together with Hiroshi Yasuda in 1988.

“consumer electronics with compact disc, broadcasting with digital radio and telecommunications with Narrowband ISDN and ASDL.” (*Id.*) The foreword recognizes that “even today, after several years and countless implementations not a single error was found in the specification.” (*Id.*) Thus, “MPEG-1 devices are now counted by the millions in all continents.” (*Id.*)

86. *Mitchell* primarily discloses an overview of MPEG standards (MPEG-1 and MPEG-2) and other concepts in details, for example, MPEG coding principles, motion compensation, motion estimation, variable quantization, rate control in MPEG, etc. For instance, the *Mitchell* textbook includes, as part of its Introduction chapter, a section entitled “Why standardize video compression?” (*Id.* at 38-39.) In it, *Mitchell* explains that “the MPEG standardization effort was initiated in 1988” as a result of the “several different industries [] converging on digital video technology.” (*Id.* at 38.) These industries included “[t]he computer industry [] looking to expand beyond its traditional text and graphic capabilities into interactive multimedia,” “[t]he consumer electronics industry [viewing] digital video as a means for improving the capabilities and interactivity of video games and entertainment media,” and “[t]he telecommunications industry [] taking advantage of the maturing compression technology to standardize teleconferencing.” (*Id.*) As *Mitchell* explains, “[w]ith the sharing of technology among these diverse industrial

interests, standardization was a natural development” as “the cost reduction implicit in having a common digital video technology for all has been a key driving force for standardization.” (*Id.*) *Mitchell* also explains that the interests of different industries converging which resulted into a common standard “lower[ed] the barriers that are usually present to the deployment of new technology.” (*Id.*)

87. *Mitchell* discloses that “[t]he basic building block of an MPEG picture is the macroblock, sketched in Figure 2.6. The macroblock consists of a 16 x 16 sample array.” (Ex. 1006 at 57 (section 2.3.3).)

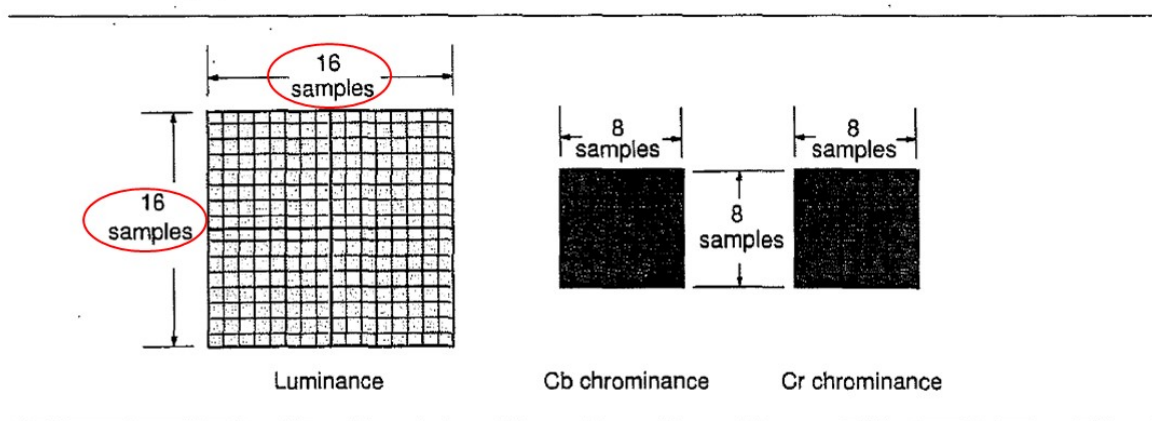


Figure 2.6: The MPEG macroblock.

(Ex. 1006 at 58 (Figure 2.6) (annotated).)

88. Further, *Mitchell* discloses MPEG-1 bit stream syntax layers (described later in detail) which includes the six layers as reflected in Figure 8.1 below: sequence layer (annotated in red), GOP layer (annotated in orange), picture layer

(annotated in green), slice layer (annotated in light blue), macroblock layer (annotated in purple), and block layer (annotated in dark blue).

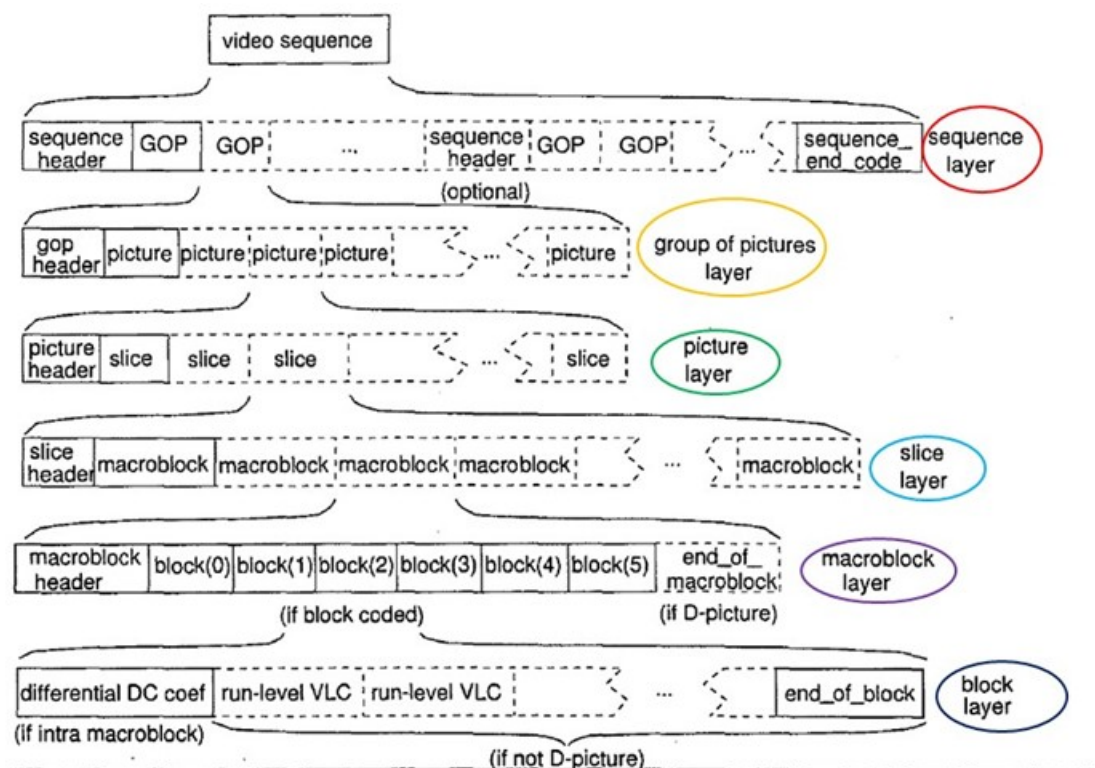


Figure 8.1: The layers of a video stream.

(*Id.* at 167 (Figure 8.1) (annotated); *see also id.* at 166 (“Figure 8.1 gives an overview of the six layers in MPEG-1 video sequences.”).)

89. In summary, *Mitchell* provides the technical background and concepts related to the MPEG standards and the MPEG standardization process to allow for a thorough understanding thereof. (*See e.g., id.* at 9.)

IX. THE PRIOR ART DISCLOSES ALL OF THE FEATURES OF CLAIMS 1-3, 5, 6, 8, 39 AND 41 OF THE '005 PATENT

A. *Hanami* Discloses the Features of Claims 1-3, 5, 6, 8, 39 and 41

90. In my opinion, *Hanami* discloses or suggests all of the features of claims 1-3, 5, 6, 8, 39 and 41. Below, I address each limitation of these claims.

1. Claim 1

a) “A method for motion coding an uncompressed digital video data stream, including the steps of:”

91. I have been asked to assume that the preamble of claim 1 is limiting. Under that assumption, it is my opinion that *Hanami* discloses the elements therein. For example, *Hanami* discloses an image coding system and corresponding techniques for coding an uncompressed digital video signal (“method for motion coding an uncompressed digital video data stream”). (Ex. 1005 at Abstract (“provided is an image coding system which can reduce the amount of codes of motion vectors with excellent picture quality”); *see also id.* at 1:5-8 (disclosing “detecting motion vectors employed for motion compensation of moving images”), 1:10-56, 6:48-55, FIGS. 1-4, 83.) As *Hanami* explains, “[a] data compression technique suppressing the data redundancy and reducing the quantity of transmission data is called highly efficient coding” and that “[s]ystems of such highly efficient coding include a predictive coding system.” (*Id.* at 1:13-18.) *Hanami* discloses that the predictive coding processing includes *motion-compensated interframe (or*

interfield) predictive coding, which in turn involves computation of motion vectors. (*Id.* at 1:18-36; *see also id.* at 1:44-2:12 (describing overall structure of an encoder for coding image data with reference to Figure 83), FIG. 83.)

92. While *Hanami* discloses these fundamental techniques with respect to a “conventional motion-compensated predictive coding method” (*id.* at 1:44-47), in my opinion, a person of ordinary skill in the art would have readily understood that *Hanami*’s disclosure (with the objective of providing a “motion vector detector”) builds on these fundamental techniques and does not repeat them in its detailed description. (*Id.* at 6:48-55, 10:55-11:11, 12:51-55, 13:23-29, FIGS. 1-4.) In other words, *Hanami* does not propose to change the “overall structure of an encoder for coding image data” as shown in Figure 83 (*id.* at 1:44-47) but instead discloses that one of its objectives is to change how certain data is processed within a macroblock layer 1100 in Figure 87 without changing the structure of the image. (*Id.* at 7:18-21 (describing one of *Hanami*’s objective is to process certain data such that “the variable code length of the motion vector of a macro block header area [1115] can be reduced”); *see also id.* at 5:49-67, 6:39-45, 12:22-37, FIG. 87.) Therefore, *Hanami*’s disclosures generally relate to methods of motion coding an uncompressed digital video data stream, as claimed.

93. Moreover, a person of ordinary skill in the art would have readily understood based on *Hanami*'s disclosure of motion-compensated predictive coding with respect to its background section (discussed above) and its detailed description section including Figures 1 and 4 (discussed below with respect to claim elements 1(b)-1(d)) that the image signals including pixel data being coded (e.g., the pixel data of a current frame and previous frame used to compute the motion vectors in the background section or the pixel data X and pixel data Y in Figure 4) are part of an uncompressed digital video data stream, as claimed. (*See, e.g., id.* at 1:10-56, 12:50-13:13, FIGS. 4, 83.) First, pixels are part of a digital image and a sequence of such digital images constitutes an uncompressed (e.g., not yet coded) digital video stream, as claimed.¹⁶ This is because the purpose of the motion-compensated predictive coding, as *Hanami* explains, is to “suppress[] the data redundancy and reduc[e] the quantity of transmission data” of “image signals having an enormous quantity of data.” (*Id.* at 1:10-32.) The concept of motion compensation (e.g., eliminating data redundancy, reducing the amount of data to be transmitted, etc.) is

¹⁶ *Hanami*'s motion-compensated coding is applied to a sequence of images comprising pixel data, i.e., a sequence of digital images, as I discuss below with reference to claim elements 1(b)-1(d).

applied to uncompressed moving images (image signals) which have an enormous quantity of data, not to already compressed data. Plus, *Hanami* is directed to a motion vector detector, and as I explained above in the Technical Background section, motion vectors are derived during the process of encoding an uncompressed signal. (See above at Section V.A.)

94. Thus, *Hanami* discloses motion coding an uncompressed digital video data stream. (See also my analysis below and accompanying citations with respect to claim elements 1(b)-1(d).)

b) [1] “comparing pixels of a first pixel array in a picture currently being coded with pixels of a plurality of second pixel arrays in at least one reference picture and”

[2] “concurrently performing motion estimation for each of a plurality of different prediction modes in order to determine which of the prediction modes is an optimum prediction mode;”

95. In my opinion, *Hanami* discloses this claim element.

96. As an initial matter, a person of ordinary skill in the art would have understood based on the knowledge of such a person and the disclosure of the '005 patent itself that the “comparing” step set forth in claim element 1(b)[1] may be performed as part of the “concurrently performing motion estimation” step set forth in claim element 1(b)[2]. That is, a person of ordinary skill in the art would have understood based on the disclosure of the '005 patent itself that “performing motion

estimation” would necessarily involve block-matching to obtain the absolute differential sums/evaluation values for a current block being coded. For example, the '005 patent discloses:

For all MPEG-2 prediction modes, the fundamental technique of motion estimation consists of comparing the current macroblock with a given 16-by-16 pixel array in the anchor picture, estimating the quality of the match according to the specified metric, and repeating this procedure for every such 16-by-16 pixel array located within the search range . . . [T]here exists a number of well-known criteria for determining the quality of the match. Among the best-known criteria are the Minimum Absolute Error (MAE), in which the metric consists of the sum of the absolute values of the differences of each of the 256 pixels in the macroblock with the corresponding pixel in the matching anchor picture macroblock; and the Minimum Square Error (MSE), in which the metric consists of the sum of the squares of the above pixel differences. In either case, the match having the smallest value of the corresponding sum is selected as the best match within the specified search range, and its horizontal and vertical positions relative to the current macroblock therefore constitute the motion vector.

(Ex. 1001 at 2:49-3:1 (emphasis added).)

97. Thus, the '005 patent discloses that the process of comparing pixels of a first pixel array in a picture currently being coded with pixels of a plurality of second pixel arrays in at least one reference picture is performed as part of the motion estimation for each of a plurality of different prediction modes. A person of ordinary skill in the art would have had this understanding as the '005 patent discloses that “the fundamental technique of motion estimation consists of comparing the current macroblock with a given 16-by-16 pixel array in the anchor picture . . . and repeating this procedure for every such 16-by-16 pixel array located within the search range” to select “the best match within the specified search range.” (*Id.*; see also above Section VIII.A.3-VIII.A.4.)

98. The block-matching disclosed by *Hanami* encompasses comparing pixels of a first pixel array in a picture currently being coded with pixels of a plurality of second pixel arrays in at least one reference picture, as claimed. In particular, as to *Hanami*'s disclosure of claim elements 1(b)[1] and 1(b)[2], *Hanami* discloses “obtaining evaluation values which indicate similarity between a current image block [e.g., ‘first pixel array’] of a prescribed size to be subjected to motion vector detection in a current screen [e.g., ‘picture currently being coded’] and respective ones of a plurality of reference image blocks [e.g., ‘plurality of second pixel arrays’] in a search area in a reference screen [e.g., ‘at least one reference picture’] related to

the current image block and are decided in accordance with respective ones of a plurality of predetermined predictive modes [e.g., ‘different prediction modes’] in parallel with each other [e.g., ‘concurrently’] for the respective reference blocks” in order to determine which of the predictive modes is an optimum prediction mode. (Ex. 1005 at 6:57-65; FIG. 4.) I discuss this in further detail below.

Claim element 1(b)[1]

99. As I described above and further discuss below, in *Hanami*’s examples, the current image block being coded is represented by a template block 1 (TB1) of a current frame 36 shown in Figures 8A-8B (reproduced below) and the reference image blocks in a search area includes the collection of macroblocks MB1-MB3 in a reference frame 37 shown in Figures 10-11 (reproduced below). (*See above* Section VIII.A.2 (noting that the search area includes, but is not limited to, macroblocks MB1-MB3).)

FIG. 8A

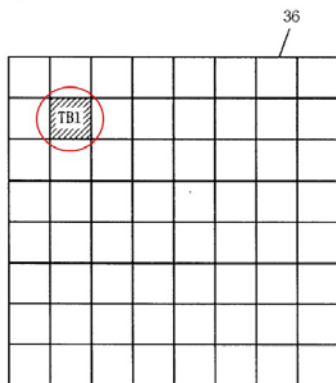
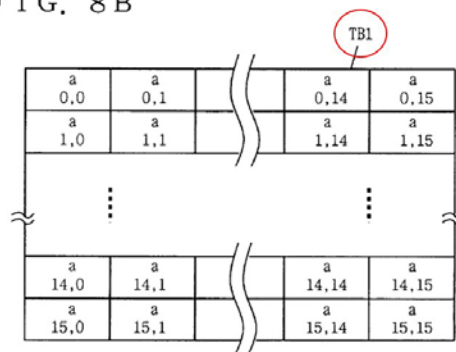


FIG. 8B



(Ex. 1005 at FIGS. 8A-8B (annotated).)

FIG. 10

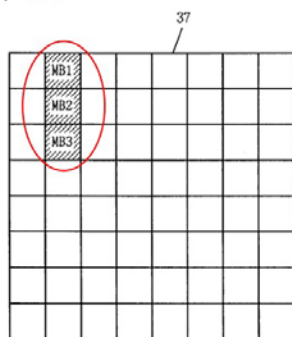
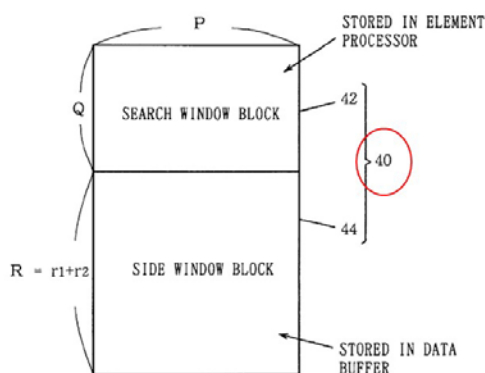


FIG. 11



(*Id.* at FIGS. 10-11 (annotated).)

100. *Hanami* discloses that the template block TB1 and the reference image blocks (e.g., MB1-MB3) included in the search area are both pixel arrays, as claimed. For example, *Hanami* discloses that template block TB1 “*is formed by the macro block* shown by slant lines in the one-frame image 36 . . . [and that] the template block TB1 includes template data $a(0, 0)$ to $a(15, 15)$ *arranged in 16 rows and 16 columns*, as shown in FIG. 8B.” (*Id.* at 15:39-47 (emphasis added); *see also id.* at 16:48-53, FIG. 15 (showing template block data 43 in the form of a 16 x 16 pixel array).) Thus, the template block TB1 discloses the claimed first pixel array. Similarly, the reference image blocks (e.g., represented by the collection of macroblocks MB1-MB3) included in the search area are also arranged in 16 rows and 16 columns. (*Id.* at 16:53-55, FIG. 15 (showing search window block data 42 in the form of a 16 x 16 pixel array).) *Hanami* discloses that the template block TB1

and reference image blocks (e.g., MB1-MB3) included in the search area both have pixels arranged in Q rows and P columns (i.e., in a form of an array). (*Id.* at 14:11-14, 15:64-16:10, FIGS. 5, 11.)

FIG. 5

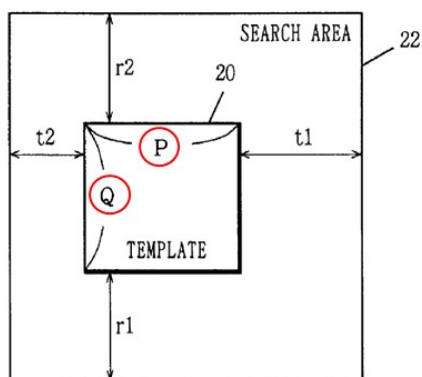
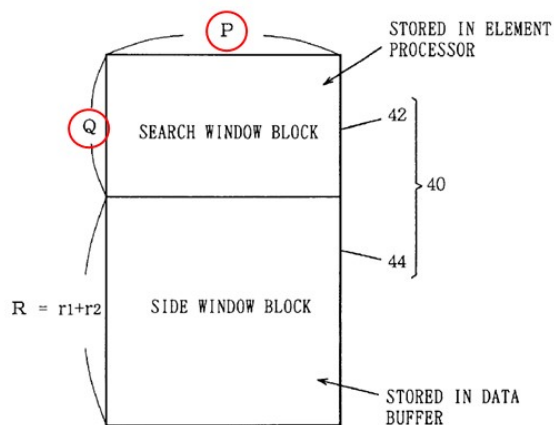
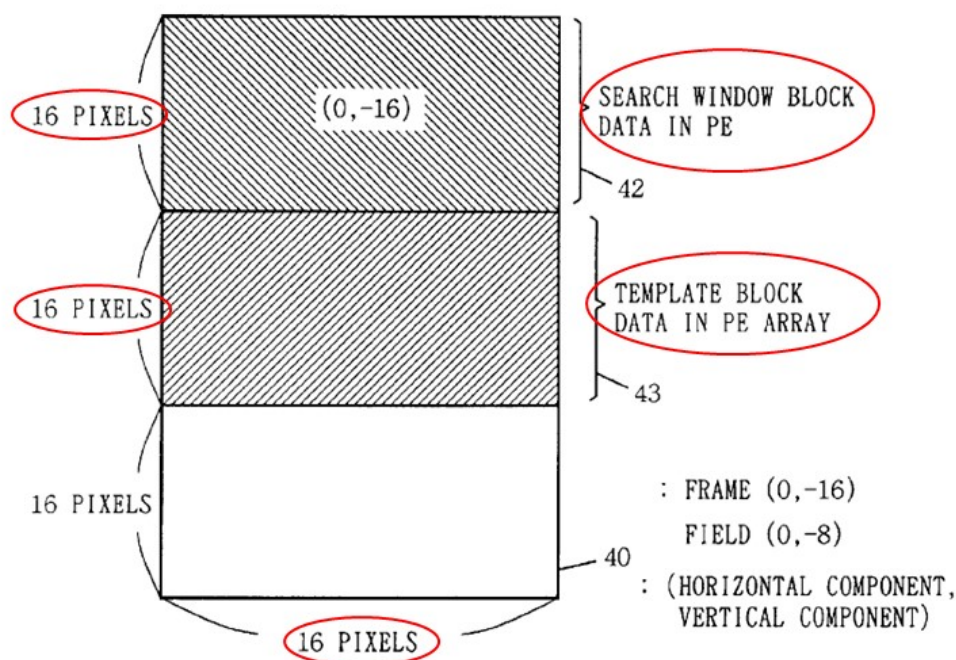


FIG. 11



(*Id.* at FIG. 5 (left) (annotated), FIG. 11 (right) (annotated).)

FIG. 15



(*Id.* at FIG. 15 (annotated).)

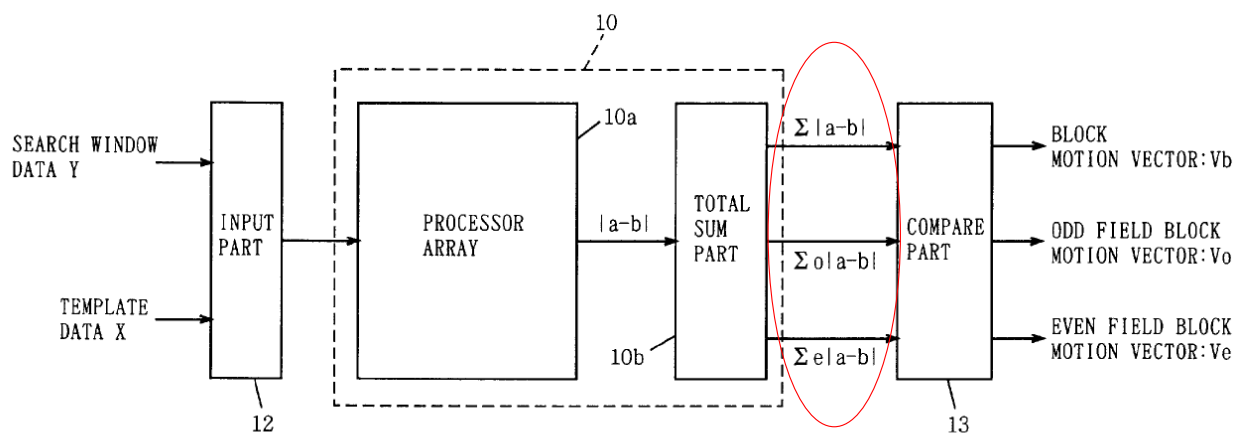
101. Coming to the comparison of pixels of the template block TB1 (of current frame 36) currently being coded and the reference image blocks (e.g., MB1-MB3) included in the search area of reference frame 37, *Hanami* discloses:

[t]he different element processors included in the processor array 10a store data of different template block pixels [e.g., “first pixel array”]. The pixel data of the reference image block in the search area [e.g., “second pixel arrays”] related to the template block are successively transmitted along a one-way direction through the element processors. The element processors provided in the processor array 10a execute operations of

the template block pixel data and pixel data of a single search window block (reference image block). Therefore, each element processor computes [e.g., “comparing”] the component of an *evaluation value* indicating the correlation between a single template block and the search window block.

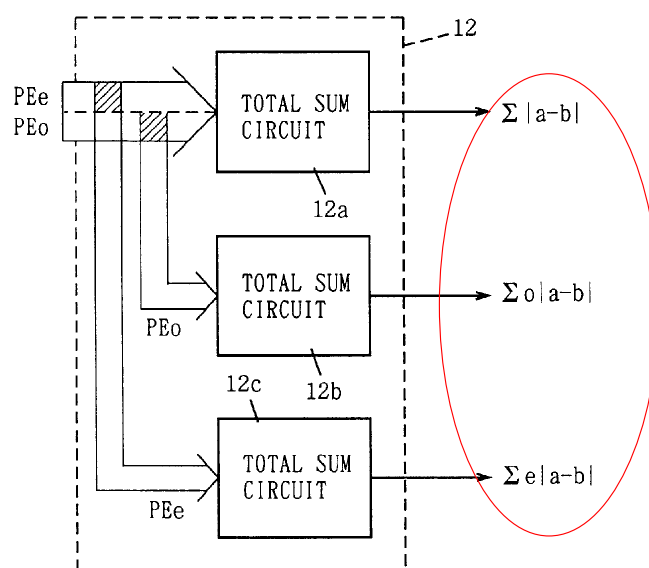
(*Id.* at 13:41-53 (emphasis added); FIG. 4; *see also id.* at 15:64-16:34; FIG. 25.)

FIG. 4



(*Id.* at FIG. 4 (annotated).)

F I G. 25



(*Id.* at FIG. 25 (annotated).)

102. That is, as shown above with respect to Figure 25 (which illustrates structure of the total sum part 10b from Figure 4), after receiving all the outputs of the element processors PE in processor array 10a in parallel with each other, “[t]he total sum circuits 12a to 12c are provided in correspondence to ***the respective predictive modes*** and ***driven in parallel with each other***, whereby ***the evaluation values $[\Sigma|a-b|, \Sigma o|a-b|$ and $\Sigma e|a-b|]$ corresponding to the respective predictive modes can be computed and outputted in parallel with each other.***” (*Id.* at 20:48-52 (emphasis added); FIG. 25; *see also id.* at 13:4-14 (explaining that evaluation values $\Sigma|a-b|$, $\Sigma o|a-b|$ and $\Sigma e|a-b|$ correspond to the frame predictive mode, the odd field predictive mode, and the even field predictive mode, respectively); 20:27-48.)

103. With the above, *Hanami* discloses “comparing pixels of a first pixel array in a picture currently being coded with pixels of a plurality of second pixel arrays in at least one reference picture,” as recited in claim element 1(b)[1], because it discloses computing evaluation values indicating a correlation between a template block TB1 (*id.* at FIGS. 8A-8B) in a current frame and reference candidate blocks (e.g., MB1-MB3) included in a search area in a reference frame (*id.* at FIGS. 10-11).

Claim element 1(b)[2]

104. In my opinion, *Hanami* additionally discloses “concurrently performing motion estimation for each of a plurality of different prediction modes in order to determine which of the prediction modes is an optimum prediction mode,” as recited in claim element 1(b)[2], because it discloses performing motion estimation in parallel (“concurrently”) for at least a frame predictive mode, odd field predictive mode, and even field predictive mode (“each of a plurality of different prediction modes”) in order to determine which of the predictive modes is an optimum predictive mode.

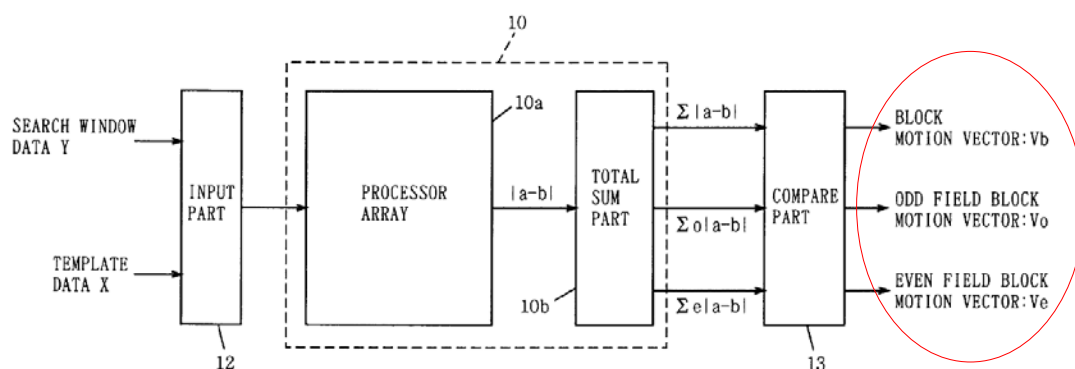
105. For example, *Hanami*’s objective is to “simultaneously detect optimum vectors efficiently coded for a plurality of predictive modes in parallel with each other” such that “[a] motion vector is produced on the basis of the optimum vector.” (*Id.* at 6:52-55, 7:13-14, 11:9-11 (disclosing that a “[m]otion vector for a single

macro block is decided on the basis of these optimum vectors [corresponding to the respective prediction modes]”), 15:38-39 (disclosing that “a motion vector is detected for each macroblock”), 24:3-8.) Thus, a person of ordinary skill in the art would have readily understood that *Hanami* performs the concurrent motion estimation to produce the evaluation values $\Sigma|a-b|$, $\Sigma_o|a-b|$ and $\Sigma_e|a-b|$ corresponding to the different predictive modes in order to determine which of the predictive modes is an optimum predictive mode, as claimed. (*See id.* at 6:1-10 (explaining that “to attain excellent coding efficiency,” it is “necessary” to perform motion detection for a plurality of predictive modes to select the optimum predictive mode and detect motion vectors in accordance with the optimum predictive mode).)

106. In other words, given that “[i]t is necessary to transmit a motion vector per pixel block” (*id.* at 3:26), once *Hanami*’s motion vector detector concurrently computes evaluation values $\Sigma|a-b|$, $\Sigma_o|a-b|$ and $\Sigma_e|a-b|$ corresponding to the frame predictive mode, the odd field predictive mode, and the even field predictive mode, respectively, these evaluation values serve as candidates for determining the optimum prediction mode. To this end, as discussed below with respect to claim element 1(c), *Hanami* discloses that its motion vector detector includes “a compare part 13 for receiving the three evaluation values $\Sigma|a-b|$, $\Sigma_o|a-b|$ and $\Sigma_e|a-b|$ obtained in the operation part 10” (*see id.* at 12:65-13:3) and “[t]he compare part 13 outputs

motion vectors (optimum vectors) decided in accordance with the respective predictive modes, i.e., a block motion vector V_b decided in accordance with the frame predictive mode, an odd field block motion vector V_o decided in accordance with the odd field predictive mode and an even field block motion vector V_e decided in accordance with the even field predictive mode in parallel with each other” (*see id.* at 13:15-23.) (*See below* Section IX.A.1(c).)

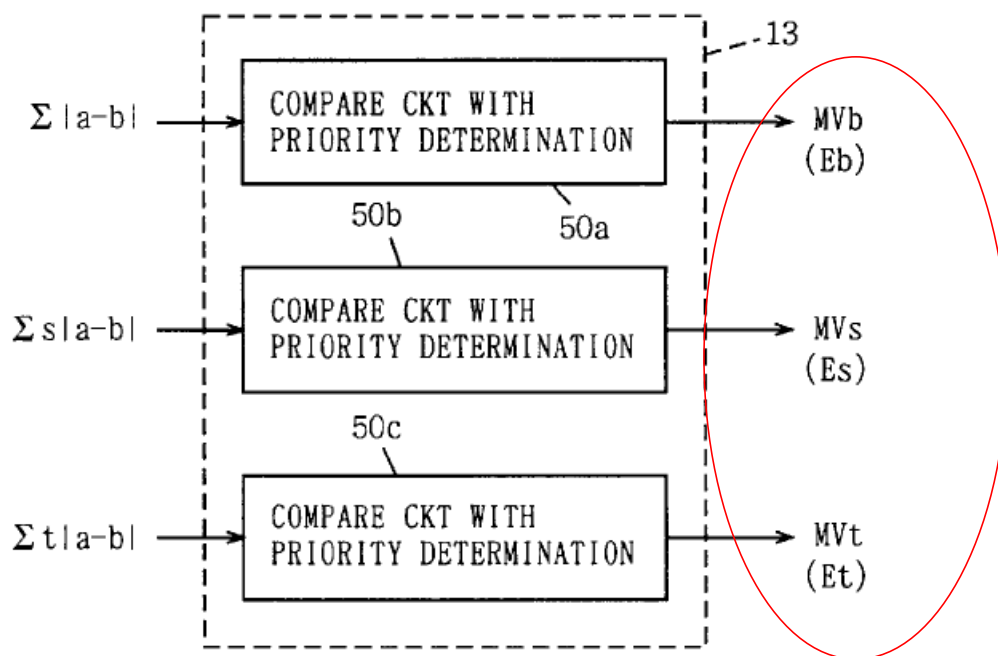
FIG. 4



(Ex. 1005 at FIG. 4 (annotated).)

107. *Hanami* discloses a detailed configuration of its compare part 13 in Figure 32. (*Id.* at 8:55-56, 23:49-24:15, FIG. 32.)

FIG. 32



(*Id.* at FIG. 32 (annotated).)

108. As an initial matter, my analysis here relies on the embodiment of compare part 13 described with respect to Figure 32 including any optional features of the compare part 13 described therein. For example, within its description of Figure 32, *Hanami* discloses that “[t]he compare part 13 evaluates the evaluation values and outputs optimum vectors MVb, MVs and MVt [annotated above in red] corresponding to the respective predictive modes,” as illustrated in Figure 32. (*Id.* at 24:1-3, FIG. 32.) *Hanami* further discloses that “why these vectors MVb, MVs and MVt are called *not motion vectors but optimum vectors* is that a structure of

performing further selection from the plurality of predictive modes and *deciding a motion vector in accordance with the predictive system having the minimum evaluation value* may be employed.” (*Id.* at 24:3-8 (emphasis added).) While this structure is described in optional terms within the Figure 32 embodiment of compare part 13, my analysis of *Hanami* here relies on the embodiment in Figure 32 of the compare part 13 including this optional structure.

109. Thus, by disclosing that its compare part 13 decides a motion vector based on the predictive mode having a minimum evaluation value, *Hanami* discloses “determin[ing] which of the prediction modes is an optimum prediction mode,” as claimed, based on the concurrent performance of motion estimation for the plural predictive modes. This is consistent with what a person of ordinary skill in the art at the time of the alleged invention would have understood as the objective of performing motion estimation (which encompasses block matching) in different prediction modes—to determine an optimum prediction mode for a current macroblock being coded. (*See, e.g.*, Ex. 1007 at 2:66-3:12 (“Several candidate predicted macroblocks may be formed for each macroblock currently being predicted by using, for example, frame based, field based and ‘dual prime’ encoding. The best prediction for a particular Current Macroblock (CM) is selected from these

candidate predicted macroblocks to be compressed and transmitted.”).) In fact, *Hanami* confirms this point as it specifically discloses

[i]n order to attain excellent coding efficiency (to reduce the quantity of coded data), it is necessary to perform motion detection in accordance with a plurality of predictive image detection systems (predictive coding systems), ***thereafter select the optimum predictive image detection system, and detect motion vectors in accordance with the optimum predictive image detection system.***¹⁷

(Ex. 1005 at 6:3-10 (emphasis added); *see also id. generally* at 6:1-37.)

110. The predictive mode having a minimum evaluation value would correspond to the claimed “optimum prediction mode” from among the frame

¹⁷ In my opinion, a person of ordinary skill in the art would have understood *Hanami*’s description of “perform[ing] motion detection in accordance with a plurality of predictive image detection systems (predictive coding systems)” to mean performing motion estimation in accordance with a plurality of predictive modes such as frame predictive mode, odd field predictive mode, and even field predictive mode. (Ex. 1005 at 6:1-37, 6:52-55, 10:57-62, 12:65-13:3, 13:4-14; 15:37-63.)

predictive mode and the even/odd-field predictive modes applied in *Hanami*'s relied-upon disclosures. (*Id.* at 24:3-8.)

1) “***Concurrently***” ***Under First Possible Construction***

111. As discussed above, I have been asked to apply my analysis under a first proposed construction for “concurrently performing motion estimation for each of a plurality of different prediction modes” as set forth in claim element 1(b)[2], which could be construed as “performing motion estimation according to each of a plurality of prediction modes during the same search operation for a given macroblock, without performing a separate search on the same macroblock for each such mode.” (*See above* Section VII.) In my opinion, *Hanami* discloses claim element 1(b)[2] under this possible construction.

112. For example, *Hanami* discloses performing motion estimation (e.g., *Hanami*'s block matching between the template block 1 (TB1) of a current frame 36 (Ex. 1005 at FIGS. 8A-8D) and reference image blocks (e.g., MB1-MB3) included a search area in reference frame 37 (e.g., *id.* at FIGS. 10-11)) according to each of a plurality of prediction modes (e.g., *Hanami*'s frame predictive mode, odd field predictive mode, and even field predictive mode) during the same search operation for a given macroblock (e.g., during the same search operation for TB1 as the total sum part 10b computes and outputs the evaluation values $[\Sigma|a-b|, \Sigma o|a-b|$ and $\Sigma e|a-$

b]] corresponding to the respective predictive modes in parallel with each other), without performing a separate search on the same macroblock (e.g., TB1) for each such mode. (*See above* Section IX.A.1(b).)

113. In particular, as I discussed above, *Hanami* repeatedly and explicitly discloses that for a given template block TB1, evaluation values [$\Sigma|a-b|$, $\Sigma o|a-b|$ and $\Sigma e|a-b|$] respectively corresponding to the frame predictive mode, odd field predictive mode, and even field predictive mode are simultaneously computed and outputted (e.g., by the total sum part 10b illustrated in Figures 4 and 25). (Ex. 1005 at 13:41-53, 20:48-52, FIGS. 4, 25; *see also id.* at 13:4-14, 15:64-16:34, 20:27-48.) (*See also above* Section IX.A.1(b).) Thus, a person of ordinary skill in the art would have understood that since these evaluation values (which are indicative of the correlation between TB1 in current frame 36 and the reference blocks in the search area in the reference frame 37) are computed in parallel for a given template block TB1, these evaluation values are being computed during the same search operation for TB1. In other words, *Hanami* does not disclose performing a separate search on the same TB1 for each of the prediction modes (for example, a separate search with respect to TB1 for the frame predictive mode, and then a separate search with respect to TB1 for the even/odd field predictive modes). Accordingly, *Hanami* discloses this claim term under its first possible construction.

2) ***“Optimum” Under First and Second Possible Constructions***

114. As discussed above, I have been asked to apply my analysis under two possible constructions for “optimum” as set forth in claim element 1(b)[2], which could be construed under the first possible construction as “best or most suitable” and under the second possible construction as at least including “the prediction mode that yields the smallest value of an error metric.” (*See above* Section VII.) In my opinion, *Hanami* discloses claim element 1(b)[2] under both these possible constructions.

115. With respect to the first possible construction, as I discussed above, *Hanami* expressly discloses that “[i]n order to attain excellent coding efficiency (to reduce the quantity of coded data), it is necessary to perform motion detection in accordance with a plurality of predictive image detection systems (predictive coding systems), ***thereafter select the optimum predictive image detection system***, and detect motion vectors in accordance with the optimum predictive image detection system.” (Ex. 1005 at 6:3-10 (emphasis added); *see also id. generally* at 6:1-37.)¹⁸ Thus, given that *Hanami* discloses selecting an “optimum” predictive mode based on the concurrent motion estimation performance for the frame predictive mode and

¹⁸ *See above* n.17.

the even/odd-field predictive modes applied in the relied-upon disclosures of *Hanami*, a person of ordinary skill in the art would have understood that the “optimum” predictive mode would have been the “best or most suitable” predictive mode under the first possible construction.

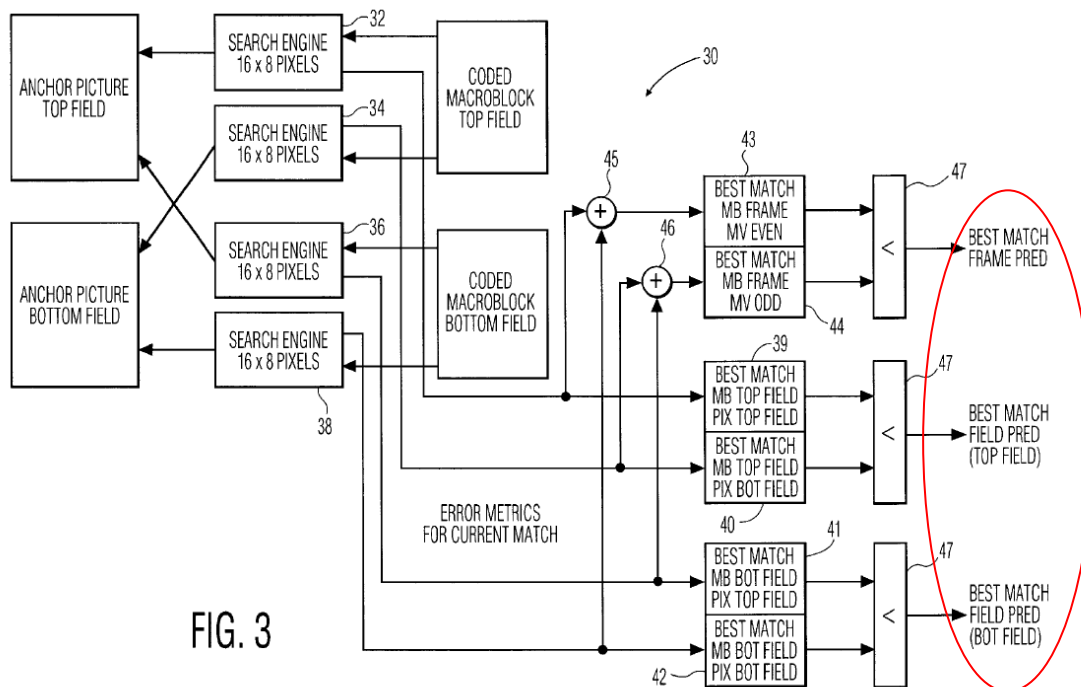
116. This is further confirmed by *Hanami*’s disclosure considering the second possible construction. In particular, as noted above, *Hanami* specifically discloses that its compare part 13 employs a structure where, from among the optimum motion vectors generated for the plurality of predictive modes, a motion vector “with the predictive system [i.e., predictive mode] having ***the minimum evaluation value***” is determined, where “[t]he absolute differential sums [representing, e.g., displacement between the block TB1 in current frame and reference blocks MB1-MB3 included in the search area in reference frame] are generally called evaluation values.” (*Id.* at 24:3-8 (emphasis added), 3:2-25.) (*See also above* Section VIII.A.3.) Thus, because *Hanami* discloses determination of a motion vector based on which predictive mode has the minimum evaluation value (“smallest value of an error metric”), *Hanami* discloses determining the prediction mode that yields the smallest value of an error metric under the second possible construction and such a predictive mode with the minimum evaluation value would have been the “best or most suitable” predictive mode under the first possible

construction as it would be the closest approximation to the current block being processed.

c) **“determining which of the second pixel arrays constitutes a best match with respect to the first pixel array for the optimum prediction mode; and,”**

117. In my opinion, *Hanami* discloses this claim element.

118. First, as I discussed above with respect to claim elements 1(b)[1] and 1(b)[2], the claimed “motion estimation” encompasses the block-matching technique. (*See above* Section IX.A.1(b).) Here, the “block-matching” technique would include “determining which of the second pixel arrays constitutes a best match with respect to the first pixel array for the optimum prediction mode,” as claimed. The ’005 patent’s disclosure is consistent with this understanding, because it does not disclose a separate step (outside of the motion estimation processing) that determines a best block match for the optimum prediction mode. (Ex. 1001 at 2:49-3:1.) In other words, the best block match for each prediction mode (including the optimum prediction mode) is identified as part of the motion estimation when the different prediction modes are concurrently applied. (*Id.*; *see also id.* at 8:49-9:27; FIG. 3 (outputs of block 47).)



(*Id.* at FIG. 3 (annotated).)

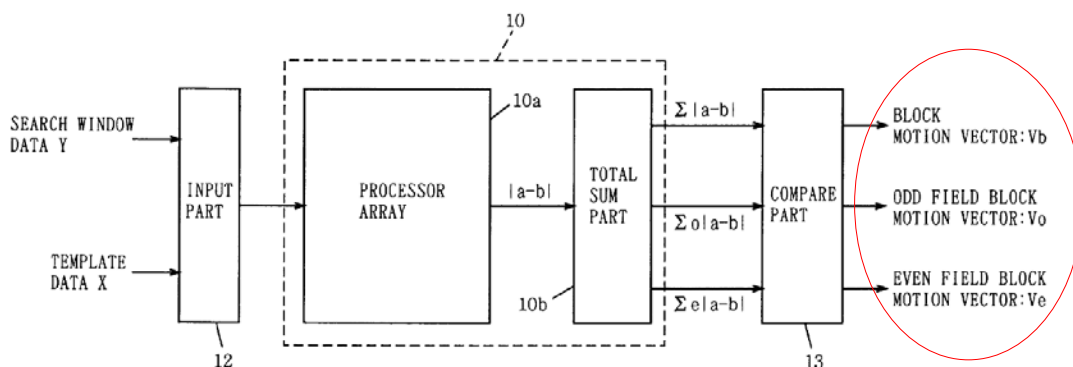
119. As to *Hanami*'s disclosure regarding claim element 1(c), *Hanami* discloses determining which of the reference candidate blocks (e.g., MB1-MB3) included in a search area in a reference frame (Ex. 1005 at FIGS. 10-11) ("second pixel arrays") constitutes a best match with respect to the template block TB1 (*id.* at FIGS. 8A-8B) ("first pixel array") for the optimum prediction mode.

120. As I discussed above with respect to claim elements 1(b)[1]-1(b)(2), *Hanami*'s objective is to "simultaneously detect optimum vectors efficiently coded for a plurality of predictive modes in parallel with each other" such that "[a] motion vector is produced on the basis of the optimum vector." (Ex. 1005 at 6:1-10, 6:52-

55, 7:13-14, 11:9-11, 15:38-39 (disclosing that “a motion vector is detected for each macroblock”), 24:3-8; *see also* Section IX.A.1(b).) And given that “[i]t is necessary to transmit a motion vector per pixel block” (Ex. 1005 at 3:27), once *Hanami*’s motion vector detector computes evaluation values $\Sigma|a-b|$, $\Sigma_o|a-b|$ and $\Sigma_e|a-b|$ corresponding to the frame predictive mode, the odd field predictive mode, and the even field predictive mode, respectively, one of these evaluation values must be considered as a best match based on the optimum prediction mode.

121. To this end, as discussed above, *Hanami* discloses that its motion vector detector includes “a compare part 13 for receiving the three evaluation values $\Sigma|a-b|$, $\Sigma_o|a-b|$ and $\Sigma_e|a-b|$ obtained in the operation part 10” (*see id.* at 12:65-13:3) and “[t]he compare part 13 outputs motion vectors (optimum vectors) decided in accordance with the respective predictive modes, i.e., a block motion vector V_b decided in accordance with the frame predictive mode, an odd field block motion vector V_o decided in accordance with the odd field predictive mode and an even field block motion vector V_e decided in accordance with the even field predictive mode in parallel with each other” (*see id.* at 13:15-23.) (*See above* Section IX.A.1(b).)

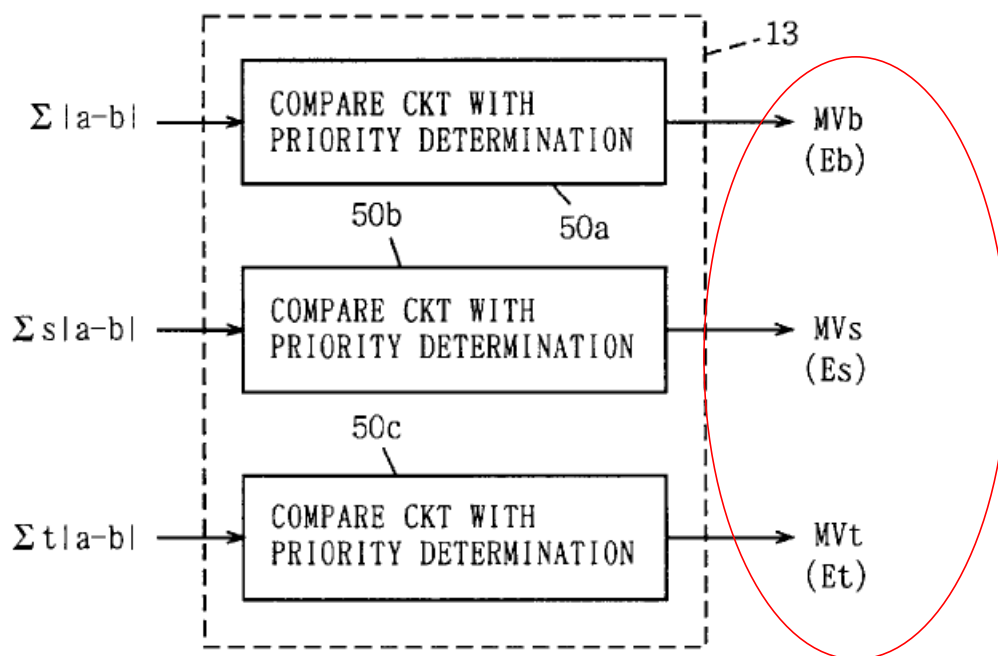
FIG. 4



(Ex. 1005 at FIG. 4 (annotated).)

122. As I discussed above with respect to claim element 1(b), *Hanami* specifically discloses that its compare part 13 employs a structure where, from among the optimum vectors generated for the plurality of predictive modes, a motion vector “with the predictive system [i.e., predictive mode] having the minimum evaluation value” is decided. (*Id.* at 24:3-8; *see also id.* at 8:55-56, 23:49-24:15.) (*See also above* Section IX.A.1(b).)

FIG. 32



(Ex. 1005 at FIG. 32 (annotated).)

123. Thus, by disclosing that a motion vector is decided based on the predictive mode having a minimum evaluation value, *Hanami* discloses “determining which of the second pixel arrays constitutes a best match with respect to the first pixel array for the optimum prediction mode,” as claimed. This is consistent with the objective of block matching as would have been understood by a person of ordinary skill in the art and confirmed by *Hanami*—as discussed above, it specifically discloses that “[i]n order to attain excellent coding efficiency (to reduce the quantity of coded data), it is necessary to perform motion detection in

accordance with a plurality of predictive image detection systems (predictive coding systems), *thereafter select the optimum predictive image detection system, and detect motion vectors in accordance with the optimum predictive image detection system.*” (*Id.* at 6:3-10 (emphasis added); *see also id. generally* at 6:1-37.) As I discussed above, the predictive mode having a minimum evaluation value (“best match”) would correspond to the claimed “optimum prediction mode” from among the frame predictive mode and the even/odd-field predictive modes applied in *Hanami*’s relied-upon disclosures. (*Id.* at 24:3-8.) (*See above* Section IX.A.1(b).)¹⁹

1) ***“Best match” Under First and Second Possible Constructions***

124. As discussed above, I have been asked to apply my analysis under two possible constructions for “best match” as set forth in claim element 1(c), which could be construed under the first possible construction as “the second pixel array within a specified search range of the reference picture having the smallest value of

¹⁹ Because claim elements 1(b) and 1(c) both encompass the same video coding process of “motion estimation” (e.g., block-matching), I cite to and rely upon similar disclosures from *Hanami* to disclose the claimed “optimum prediction mode” in claim element 1(b) and “best match” in claim element 1(c). (*See above* Section IX.A.1(b).)

an error metric” and under the second possible construction as at least including the array included in the search “having the smallest value of an error metric.” (*See above* Section VII.) *Hanami* discloses claim element 1(c) under both these possible constructions.

125. Under either construction, as discussed above, *Hanami* discloses that its compare part 13 employs a structure where, from among the optimum motion vectors generated for the plurality of predictive modes, a motion vector “with the predictive system [i.e., predictive mode] having *the minimum evaluation value*” is determined, where “[t]he absolute differential sums [representing, e.g., displacement between the block TB1 in current frame and reference blocks (e.g., MB1-MB3) included in a search area in reference frame] are generally called evaluation values.” (Ex. 1005 at 24:3-8 (emphasis added), 3:2-25.) (*See also above* Sections VIII.A.3-4, IX.A.1(b).) Thus, because *Hanami* discloses determination of a motion vector based on which predictive mode has the minimum evaluation value, a person of ordinary skill in the art would have understood that *Hanami* discloses determining which one of the reference blocks (e.g., MB1-MB3) included in the search area in reference frame 37 (“the second pixel array within a specified search range of the reference picture”) has the minimum evaluation value (“having the smallest value of

an error metric”) as set forth in the first and second possible constructions. (*See above* Section IX.A.1(b)(2).)²⁰

d) **“generating a motion vector for the first pixel array in response to the determining step.”**

126. In my opinion, *Hanami* discloses this claim element, including for reasons similar to those discussed above with respect to claim element 1(c).

127. In particular, as discussed above with respect to Figures 4 and 32, the compare circuit 13 outputs optimum vectors (block motion vector V_b corresponding to frame predictive mode, odd field block motion vector V_o corresponding to the odd field predictive mode, and even field block motion vector V_e corresponding to the even field predictive mode) based on receiving the evaluation values $\Sigma|a-b|$, $\Sigma_o|a-b|$ and $\Sigma_e|a-b|$ obtained from the operation part 10. (*See above* Section IX.A.1(c).)

128. *Hanami* discloses with respect to its overall motion vector detection structure that after the optimum vectors are generated based on the evaluation values computed for the different predictive modes, a “[m]otion vector for a single macro block [i.e., the template block] is decided on the basis of these optimum vectors.” (Ex. 1005 at 11:9-11; *see also id.* at 3:26, 3:39-40, 15:38-39, 24:3-8.) *Hanami* specifically discloses that its compare part 13 employs a structure where, from

²⁰ *See above* n.19.

among the optimum motion vectors generated for the plurality of predictive modes, a motion vector “with the predictive system [i.e., predictive mode] having the minimum evaluation value” is determined. (*Id.* at 24:3-8, FIG. 32.) (*See also above* Section IX.A.1(b).) Thus, *Hanami* discloses generating a motion vector for the template block TB1 (e.g., “first pixel array”) in response to determining the best block match for the optimum prediction mode (e.g., the prediction mode having a minimum evaluation value).

129. A person of ordinary skill in the art would have understood that *Hanami*’s motion vector detector chooses one of the optimum vectors generated by the compare part 13 as a motion vector for the template block TB1 based on, e.g., the minimum absolute value of the corresponding evaluation values of the motion vectors. This is confirmed by the above-noted disclosure of *Hanami* with respect to Figure 32 (*see, e.g.*, Ex. 1005 at 24:3-8) and also because, as *Hanami* acknowledges, “[i]t is necessary to transmit a motion vector per pixel block.” (*Id.* at 3:26; *see also id.* at 3:39-40 (“Motion vector detection is executed in the units of the blocks.”).) Thus, for each template block being coded, a motion vector must be chosen for transmission to a decoder (after the motion vector is encoded) based on whether it corresponds to the best block match for the optimum prediction mode.

130. This is consistent with the '005 patent's disclosure's description of a "motion vector" and how the motion vector for the best match is output based on the optimum prediction mode. For instance, the '005 patent's disclosure describes a motion vector in the context of the "fundamental technique of motion estimation" (Ex. 1001 at 2:49-50), stating that "the match having the smallest value of the corresponding sum is selected as the best match within the specified search range, and its horizontal and vertical positions relative to the current macroblock therefore constitute the motion vector." (*Id.* at 2:64-3:1.)²¹ The '005 patent further states that once the best match results for the different prediction modes are compared to identify the optimum prediction mode, "[t]he corresponding motion vector for the best match produced by the selected prediction mode is then output." (*Id.* at 9:28-35.)

131. As such, a person of ordinary skill in the art would have understood in light of the '005 patent's disclosure that the claimed "generating a motion vector for the first pixel array in response to the determining step" as discussed in claim

²¹ Similarly, *Hanami* discloses that "[t]he displacement between the block of interest and the most approximate block in the preceding frame is called a motion vector." (Ex. 1005 at 3:6-8.)

element 1(d) encompasses outputting a motion vector chosen or selected from amongst the vectors corresponding to the best match results for the different prediction modes. In my opinion, this is what *Hanami* discloses—i.e., as discussed above, *Hanami* discloses that, after the optimum vectors are generated based on the evaluation values computed for the different predictive modes, a “[m]otion vector for a single macro block [i.e., the template block] is decided on the basis of these optimum vectors.” (Ex. 1005 at 11:9-11; *see also id.* at 3:26, 3:39-40, 15:38-39, 24:3-8.)

132. *Hanami* goes to great pains to reiterate that the vectors MVb, MVs and MVt output from its compare part 13 in Figure 32 are not called “motion vectors” but instead “optimum vectors” because an additional structure for “performing further selection from the plurality of predictive modes and ***deciding a motion vector*** in accordance with the predictive system having the minimum evaluation value may be employed.” (*Id.* at 24:3-8 (emphasis added).) Therefore, by disclosing that “a motion vector” is decided for the template block currently being coded based on the predictive mode having the minimum evaluation value, *Hanami* discloses generating a motion vector for the template block TB1 (“first pixel array”) in response to determining the best block match for the optimum prediction mode.

133. As I discussed above, it was suggested during prosecution that “generating a motion vector,” as claimed, is limited to newly creating a motion vector after determining the best match corresponding to the optimum prediction mode. (*See above* Section VI.B.) In my opinion, even if this claim element was interpreted this way, *Hanami* discloses this feature as it discloses “***performing further selection*** from the plurality of predictive modes and ***deciding a motion vector*** in accordance with the predictive system having the minimum evaluation value.” (Ex. 1005 at 24:3-8; *see also id.* at 11:9-11, 15:38-39.) That is, in my opinion, a person of ordinary skill in the art would have understood that until the “further selection” from the plural predictive modes is performed, no “motion vector” for the template block TB1 currently being coded is decided and thus no specific “motion vector” for the template block TB1 exists at that time since each of the optimum vectors would be a candidate vector. Once the “further selection” is made based on the minimum evaluation value from amongst the different predictive modes, then “a motion vector” is decided, i.e., a “motion vector” is generated for the template block TB1.

2. Claim 2

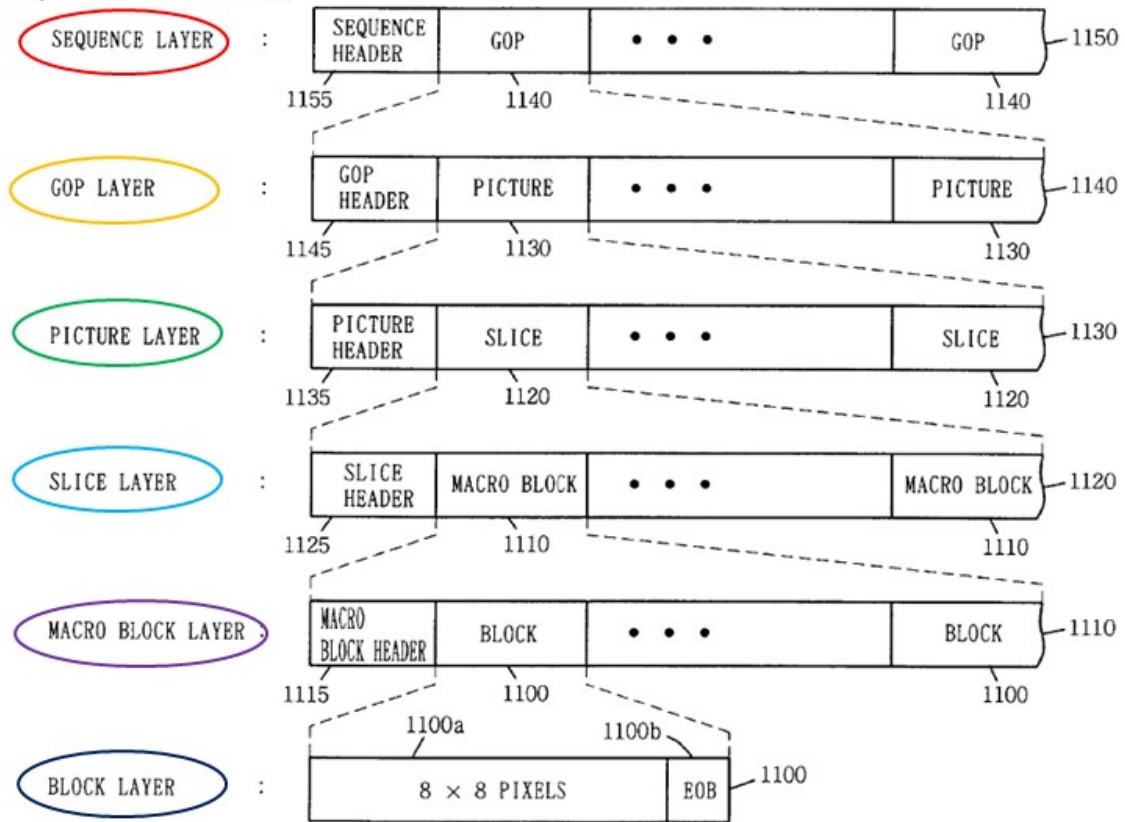
a) **“The method as set forth in claim 1, wherein the first and second pixel arrays each have a size and structure defined by an MPEG standard.”**

134. In my opinion, *Hanami* discloses this claim.

135. For example, *Hanami* discloses that the template block TB1 (*id.* at FIGS. 8A-8B) (“first pixel array”) and the candidate reference blocks (e.g., MB1-MB3) (FIGS. 10-11) included in the search area (“second pixel arrays”) each have a size and structure defined by an MPEG standard.

136. In particular, *Hanami* discloses the format of its image signals (including the template block in a current frame and candidate reference blocks within a search area in a reference frame) by way of its background with respect to Figure 87.

FIG. 87 PRIOR ART



(Ex. 1005 at FIG. 87 (annotated).)

137. *Hanami* discloses that “FIG. 87 schematically illustrates *the structure (syntax)* of image data coded by the video multiplex coding circuit 914 shown in FIG. 83.”²² (*Id.* at 4:1-3 (emphasis added), FIG. 87.) Specifically, with reference to

²² As I discussed above with respect to claim element 1(a), a person of ordinary skill in the art would have understood that *Hanami* does not propose to change the

annotated Figure 87 above, *Hanami* discloses that “the bit stream is split into a plurality of layers, i.e., a sequence layer [annotated in red], a GOP (group of picture) layer [annotated in orange], a picture layer [annotated in green], a slice layer [annotated in light blue], a macro block layer [annotated in purple] and a block layer [annotated in dark blue] in order from the uppermost layer.” (*Id.* at 4:7-10.) And *Hanami* discloses with respect to its macro block 1110 (e.g., template block TB1 and reference blocks such as macroblocks MB1-MB3 included in the search area) that it “includes a prescribed number of (six) blocks 1100” and “a macro block header 1115.”²³ (*Id.* at 4:17-23.)

“overall structure of an encoder for coding image data” as shown in Figure 83. (Ex. 1005 at 1:44-47; FIG. 83; *see also above* Section IX.A.1(a).) Instead, *Hanami* proposes to change how certain data is processed within the macroblock layer 1100 in Figure 87 without changing the structure of the image. (Ex. 1005. at 7:18-21 (describing one of *Hanami*’s objectives is to process certain data such that “the variable code length of the motion vector of a macro block header area [1115] can be reduced.”); *see also id.* at 5:49-67; 6:39-45; 12:22-37, FIG. 87.)

²³ *Hanami*’s template block TB1 in current frame 36 and reference candidate blocks (e.g., MB1-MB3) in reference frame 37 are all macroblocks. (Ex. 1005 at 15:39-41

138. This is the exact same structure of MPEG, as confirmed by contemporaneous MPEG-related literature at the time. For example, the authoritative and well-known MPEG textbook at the time of the alleged invention entitled “MPEG Video Compression Standard” by *Mitchell et al.* (Ex. 1006) (“*Mitchell*”) (which I refer to above in my discussion of the Technical Background in Section V) discloses that the **MPEG-1** bit stream syntax layers include the same six layers as reflected in Figure 8.1 below: sequence layer (annotated in red), GOP layer (annotated in orange), picture layer (annotated in green), slice layer (annotated in light blue), macroblock layer (annotated in purple), and block layer (annotated in dark blue).²⁴ (*Id.* at 166-168, FIG. 8.1.) And just like *Hanami*, *Mitchell* discloses

(“template block TB1 for motion vector detection is formed by the macro block shown by slant lines in the one-frame image 36 [in Figure 8A].”), 15:64-66 (“In order to search for a reference image for detecting a predictive image, the processor array 10a stores three macroblocks MB1, MB2 and MB3 in a preceding frame image 37, as shown in FIG. 10.”).)

²⁴ In this portion of my analysis, I rely solely on the teachings of *Hanami* for disclosing the features of claim 2. I discuss *Mitchell* (Ex. 1006) to demonstrate the knowledge of a person of ordinary skill in the art at the time of the alleged invention

that “[t]he macroblock is a group of six 8x8 DCT blocks” and “starts with a macroblock header.” (*Id.* at 167-168.)

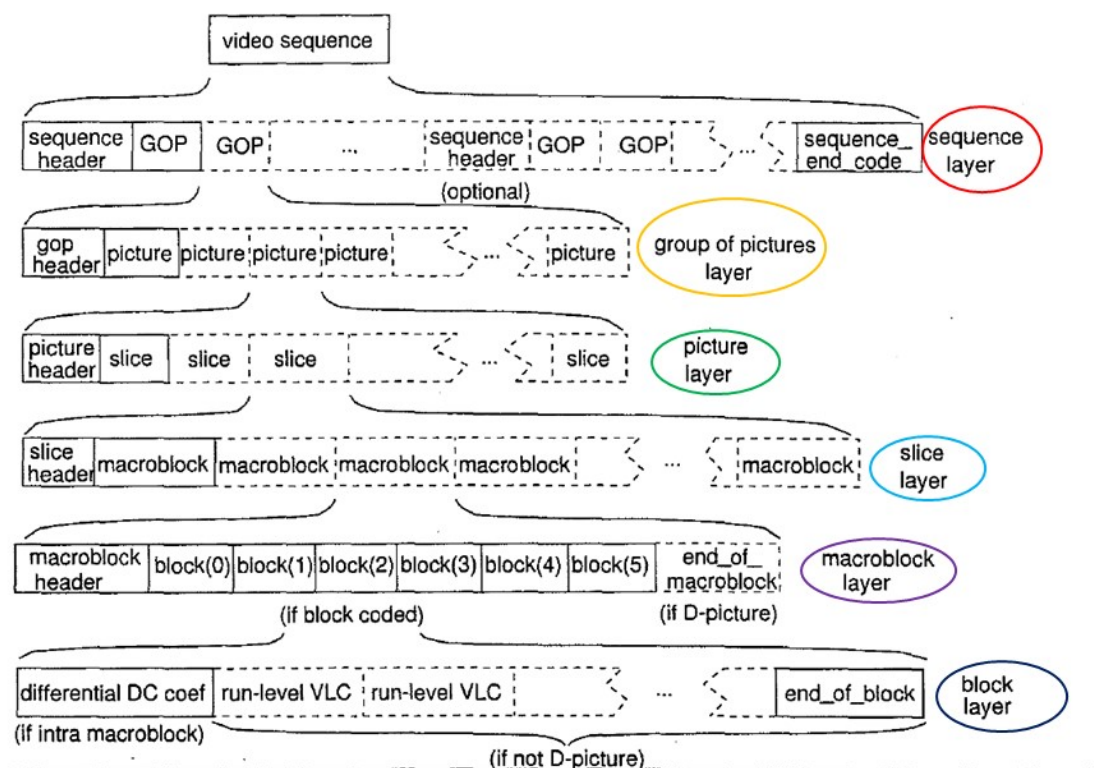


Figure 8.1: The layers of a video stream.

(*Id.* at 167 (Figure 8.1) (annotated); *see also id.* at 166 (“Figure 8.1 gives an overview of the six layers in MPEG-1 video sequences.”).)

for the '005 patent. Given such knowledge, such a skilled person, in my opinion, would have understood based on the disclosures of *Hanami* that its template block TB1 and reference candidate blocks (e.g., MB1-MB3) each have a structure and size defined by an MPEG standard.

139. In fact, a side-by-side comparison of the known MPEG structure of video sequences and that disclosed in *Hanami*, as shown below, readily confirms that the video sequence structure in *Hanami* is defined by an MPEG Standard.

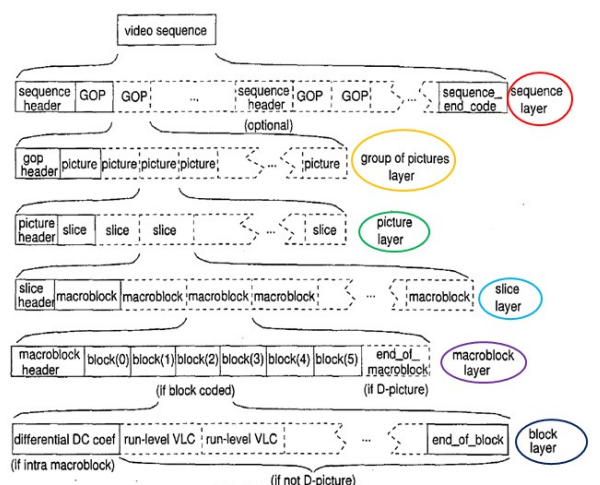
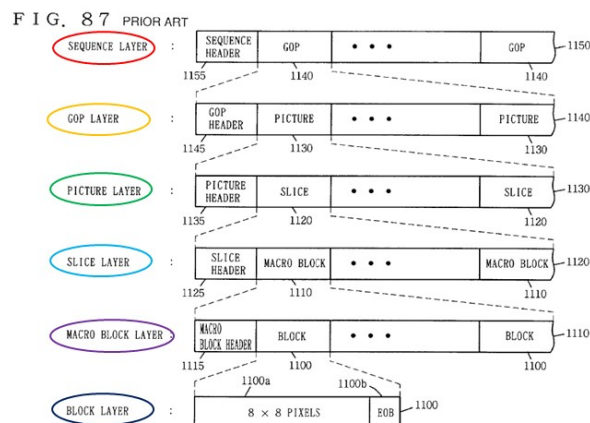


Figure 8.1: The layers of a video stream.



(*Id.* at 167 (Figure 8.1) (annotated) (left); Ex. 1005 at FIG. 87 (annotated) (right).)

140. As to the “size” of *Hanami*’s template and reference candidate blocks (e.g., the first pixel array or the second pixel arrays), *Hanami* discloses that its template block TB1 (FIGS. 8A-8B) (“first pixel array”) and reference image blocks such as macroblocks MB1-MB3 included in the search area (FIGS. 10-11) (“second pixel arrays”) are of size 16x16, as I discussed above with respect to claim element 1(b). (*See above* Section IX.A.1(b).) For example, *Hanami* discloses:

It is assumed that a *template block TB1 for motion vector detection is formed by the macro block* shown by slant lines in the one-frame image 36 . . . It is assumed that the

template block TB1 includes template data $a(0, 0)$ to $a(15, 15)$ **arranged in 16 rows and 16 columns**, as shown in FIG. 8B.

(Ex. 1005 at 15:39-47 (emphasis added), FIGS. 8A-8B.) Similarly, as I discussed above with respect to claim element 1(b), the reference image blocks (e.g., represented by the collection of macroblocks MB1-MB3) included in a search area are also arranged in 16 rows and 16 columns. (*Id.* at 16:53-55, FIG. 15 (showing search window block data 42 in the form of a 16 x 16 pixel array).)

141. Again, this is the same exact size of blocks in MPEG, as confirmed by the *Mitchell* textbook. In particular, *Mitchell* discloses that “[t]he basic building block of an MPEG picture is the macroblock, sketched in Figure 2.6. **The macroblock consists of a 16 x 16 sample array.**” (Ex. 1006 at 57 (section 2.3.3) (emphasis added).)

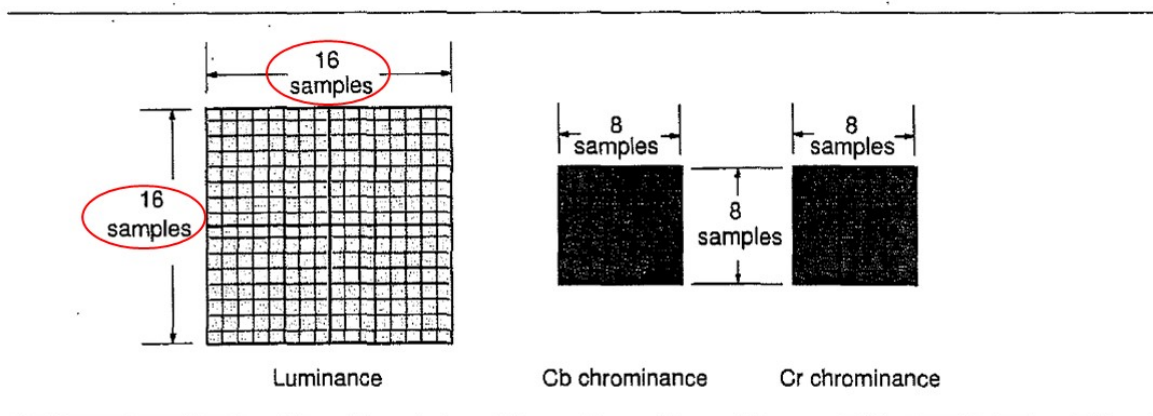


Figure 2.6: The MPEG macroblock.

(Ex. 1006 at 58 (Figure 2.6) (annotated).)

142. This understanding of the size and structure of the template block TB1 and the candidate reference blocks (e.g., MB1-MB3) included in a search area is consistent with the '005 patent. For instance, the '005 patent discloses that “[t]he MPEG picture and macroblock structure is diagrammatically illustrated in FIG. 1,” and Figure 1 shows a 16x16 macroblock just like the macroblocks (template block TB1 and the candidate reference blocks (e.g., MB1-MB3)) in *Hanami*. (Ex. 1001 at 2:26-27, FIG. 1.)

143. Thus, a person of ordinary skill in the art would have readily understood that the template block TB1 (“first pixel array”) and the candidate reference blocks (e.g., MB1-MB3) included in the search area (“second pixel arrays”) in *Hanami* each have a size and structure defined by an MPEG standard, as claimed.

3. Claim 3

a) **“The method as set forth in claim 1, wherein the method is implemented using a motion estimation search engine of a digital video encoder.”**

144. In my opinion, *Hanami* discloses this claim.

145. As an initial matter, the '005 patent discloses in relation to this feature that:

[f]or all MPEG-2 prediction modes, the fundamental technique of *motion estimation consists of comparing the*

current macroblock with a given 16-by-16 pixel array in the anchor picture, estimating the quality of the match according to the specified metric, and repeating this procedure for every such 16-by-16 pixel array located within the search range. ***The hardware or software apparatus that performs this search is usually termed the “search engine.”***

(Ex. 1001 at 2:49-56 (emphasis added).) Thus, in my opinion, any hardware or software in *Hanami* that performs the method set forth in claim 1 discloses the claimed motion estimation search engine of a digital video encoder.

146. *Hanami* discloses that its method is implemented using a source coding circuit 912 (“motion estimation search engine”) of a digital video encoder shown in Figure 83, as discussed above with respect to claim element 1(a) and further discussed below. (*See above* Section IX.A.1(a).)

147. For example, *Hanami* discloses an “overall structure of an encoder for coding image data” as shown in Figure 83.²⁵ (Ex. 1005 at 1:44-47; FIG. 83.) The encoder includes:

source coding circuit 912 [which] performs an orthogonal transformation such as discrete cosine transformation

²⁵ *See also above* Section IX.A.1(a); *see also above* n.22.

(DCT), for example, on the supplied signal, *performs motion compensation on the input signal*, and quantizes orthogonal-transformed image data.

(*Id.* at 1:61-65 (emphasis added); *see also id.* at 2:13-67 (describing additional details of the source coding circuit 912 with reference to Figure 84); FIGS. 83-84.)

FIG. 83 PRIOR ART

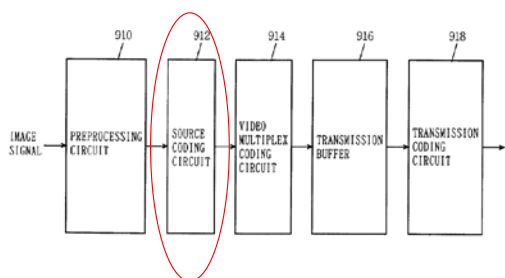
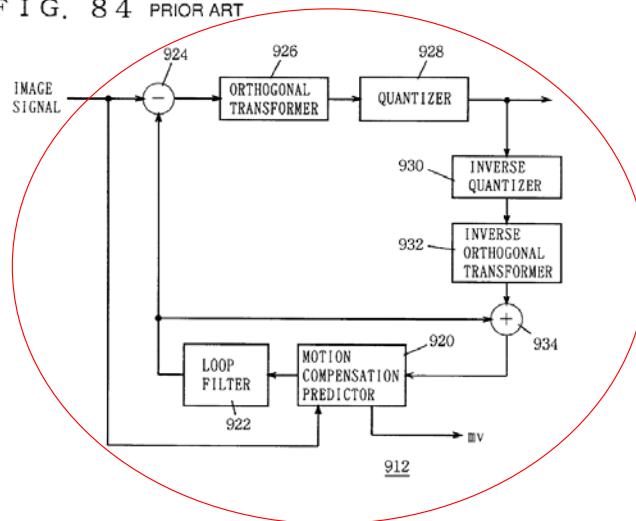


FIG. 84 PRIOR ART



(*Id.* at FIG. 83 (left) (annotated), FIG. 84 (right) (annotated).)

148. Specifically, the source coding circuit 912 includes a motion compensation predictor 920 which, among other things, detects motion vectors. (*Id.* at 2:27-33.) As I discussed above, *Hanami* explains that the generation of motion vectors involves block matching. (*Id.* at 2:65-67; *see also above* Section IX.A.1(b).) In other words, the source coding circuit 912 is the hardware or software that performs the search (block-matching) and thus is a “motion estimation search engine,” as claimed. (Ex. 1001 at 2:49-56.)

4. Claim 5

a) **“The method as set forth in claim 1, wherein the motion coding is performed in accordance with an MPEG standard.”**

149. In my opinion, *Hanami* discloses this claim for at least reasons similar to those discussed above with respect to claim 2. (*See above* Section IX.A.2.) For example, as discussed above with respect to claim 2, *Hanami* discloses a video sequence structure that is defined by the MPEG Standard. Thus, a person of ordinary skill in the art would have understood that the motion coding in *Hanami* is in accordance with an MPEG standard since its video sequence structure is compliant with the MPEG Standard.

150. In my opinion, such an understanding of a person of ordinary skill in the art is further confirmed by *Hanami*'s disclosure related to motion coding of motion vectors which is identical to the requirements imposed by MPEG-2 for motion coding of motion vectors. For example, *Hanami* discloses that Figure 3 shows “a list of variable-length codes of the variable-length motion vector components MHC [horizontal component] and MVC [vertical component].” (Ex. 1005 at 12:22-23, FIG. 3; *see also id.* at 11:66-12:5.) The MPEG-2 Part 2 standard (Ex. 1010) discloses Table B.10 which includes variable length codes for motion

coding the motion vectors. (Ex. 1010 at 141 (Table B.10).)²⁶ Table B.10 in the MPEG-2 standard is identical to Figure 3 of *Hanami*. (Compare Ex. 1005 at FIG. 3 with Ex. 1010 at 141 (Table B.10).)

²⁶ Exhibit 1010 is a copy of the MPEG-2 Part 2 standard published in 1996, which is officially known as the ISO/IEC 13818-2 International Standard. (Ex. 1010 at 1 (showing official name), 166 (section D.9 disclosing that “MPEG-2 refers to this Specification,” *i.e.*, the Specification of the ISO/IEC 13818-2 International Standard).) I am personally aware that the MPEG-2 Part 2 standard was made widely available to the public at the time of its release in 1996, and there were no confidentiality restrictions regarding further distribution and dissemination of the standard. *Mitchell* confirms that the MPEG-2 Part 2 standard (ISO/IEC 13818-2) was published in 1996. (Ex. 1006 at 422 (Table 18.3 noting that Part 2 of MPEG-2 published in 1996); *see also id.* at 205 (MPEG-2 Overview chapter noting in section 9.3 the reference to the ISO/IEC 13818-2 standard).)

FIG. 3

MHC, MVC

VARIABLE-LENGTH CODE	DECIMAL CODE
0000 0011 001	-16
0000 0011 011	-15
0000 0011 101	-14
0000 0011 111	-13
0000 0100 001	-12
0000 0100 011	-11
0000 0100 11	-10
0000 0101 01	-9
0000 0101 11	-8
0000 0111	-7
0000 1001	-6
0000 1011	-5
0000 111	-4
0001 1	-3
0011	-2
011	-1
1	0
010	1
0010	2
0001 0	3
0000 110	4
0000 1010	5
0000 1000	6
0000 0110	7
0000 0101 10	8
0000 0101 00	9
0000 0100 10	10
0000 0100 010	11
0000 0100 000	12
0000 0011 110	13
0000 0011 100	14
0000 0011 010	15
0000 0011 000	16

Table B.10 – Variable length codes for motion_code

Variable length code	motion_code[r][s][t]
0000 0011 001	-16
0000 0011 011	-15
0000 0011 101	-14
0000 0011 111	-13
0000 0100 001	-12
0000 0100 011	-11
0000 0100 11	-10
0000 0101 01	-9
0000 0101 11	-8
0000 0111	-7
0000 1001	-6
0000 1011	-5
0000 111	-4
0001 1	-3
0011	-2
011	-1
1	0
010	1
0010	2
0001 0	3
0000 110	4
0000 1010	5
0000 1000	6
0000 0110	7
0000 0101 10	8
0000 0101 00	9
0000 0100 10	10
0000 0100 010	11
0000 0100 000	12
0000 0011 110	13
0000 0011 100	14
0000 0011 010	15
0000 0011 000	16

(Ex. 1005 at FIG. 5 (left); Ex. 1010 at 141 (Table B.10) (right) (annotated).)

151. Thus, in my opinion, *Hanami* discloses the motion coding is performed in accordance with an MPEG standard.

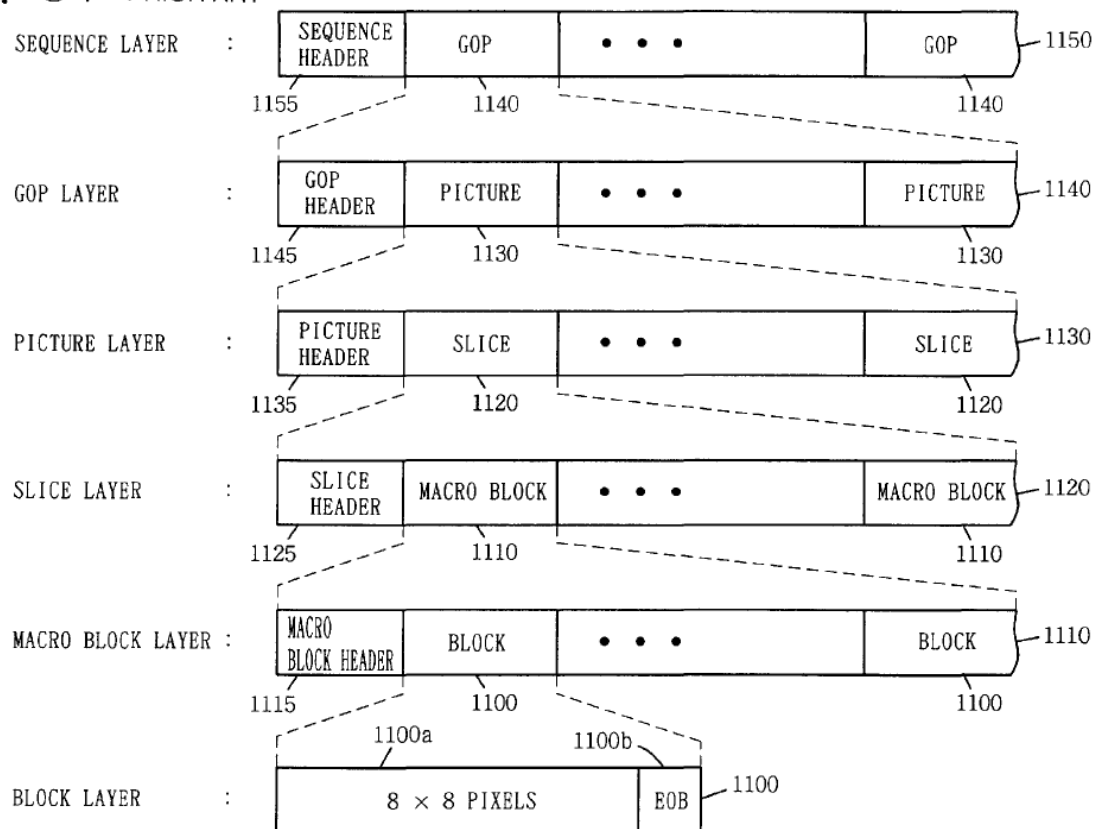
5. Claim 6

a) **“The method as set forth in claim 5, further comprising the initial step of providing information identifying a picture type of the first pixel array and using this information in the comparing step.”**

152. In my opinion, *Hanami* discloses this claim.

153. For example, *Hanami* discloses that its picture layer 1130 in Figure 87 includes a picture header 1135 which stores “information indicating the type (I picture, P picture or the like) of the picture 1130.” (Ex. 1005 at 4:32-36.)

FIG. 87 PRIOR ART



(*Id.* at FIG. 87.)

154. *Hanami* also explains that there are several different picture types.

(a) **I picture**: This is an image coded with only information closed in the image. Namely, pixel data of the I picture is coded with no difference computing.

(b) **P picture**: This is an image subjected to interframe or interfield predictive coding. A predictive image (reference image) employed for the P picture is an already coded I or P picture temporally preceding in input order. In general, it is possible to select a more efficient one of a method of coding the difference between the P picture and the motion-compensated predictive image or a method of coding the P picture with no difference computing in the units of macro blocks.

(c) **B picture** (bidirectional predictive coded image): A predictive image employed for the B picture is selected from (i) an already decoded I or P picture temporally preceding the B picture, (ii) an already decoded I or P picture temporally subsequent to the B picture, and (iii) an interpolated image produced by the pictures (i) and (ii). It is possible to select the most efficient one of methods of coding differences between the B pictures and the three types of coding of the differences after motion compensation and coding with no difference computing.

(*Id.* at 4:38-63 (emphasis added).)

155. As explained by *Hanami*, the picture type (e.g., I picture, P picture, or B picture) dictates how the image is coded (e.g., an I picture is coded with only information closed in the image, a P picture is coded based on an already coded I or P picture temporally preceding the current P picture, and a B picture is coded based on an already coded I or P picture temporally preceding the current P picture, an already coded I or P picture temporally subsequent to the current P picture, or an interpolated image produced by the temporally preceding I/P picture and the temporally subsequent I/P picture.) (*Id.* at 4:38-63.) Thus, *Hanami* discloses the initial step of providing information identifying a picture type of the template block (e.g., the picture type of the current frame 36 within which the template block TB1 resides in Figure 8A), and using this information during the block-matching (“comparing step”), e.g., the block-matching will depend on whether the current picture is coded based on only temporally preceding I/P picture(s) or only temporally sequential I/P picture(s), or both.

6. Claim 8

a) **“The method as set forth in claim 1, wherein the different prediction modes are: frame prediction mode for frame pictures; and, field prediction mode for frame pictures.”**

156. In my opinion, *Hanami* discloses this claim.

157. For example, *Hanami* discloses that the different prediction modes are frame predictive mode (“frame prediction mode”) for frame pictures and even/odd-field prediction modes (“field prediction mode”) for frame pictures.

158. In particular, *Hanami* discloses with respect to the motion vector detector illustrated in Figure 4 that it is “adapted to code pixel data in ***units of frames.***” (Ex. 1005 at 12:54-55 (emphasis added); FIG. 4.) Thus, the predictive modes described with reference to *Hanami*’s Figure 4 are for frame pictures. Moreover, as discussed above with respect to claim element 1(b)[2], the motion vector detector of Figure 4 calculates evaluation values $\Sigma|a-b|$, $\Sigma_o|a-b|$ and $\Sigma_e|a-b|$, where “[t]he evaluation value $\Sigma|a-b|$ corresponds to that of a ***frame predictive mode,***” “[t]he evaluation value $\Sigma_o|a-b|$ corresponds to that in accordance with an ***odd field predictive mode,***” and “[t]he evaluation value $\Sigma_e|a-b|$ is that computed in accordance with an ***even field predictive mode.***” (*Id.* at 13:4-14 (emphasis added); FIG. 4.) That is, the different prediction modes in *Hanami*’s relied upon disclosures are a frame predictive mode and odd/even field predictive modes for frame pictures, as claimed, since Figure 4’s motion vector detector is configured to code pixel data in units of frames.

7. Claim 39

a) **“A motion estimation search system that concurrently performs motion estimation using each of a plurality of different motion estimation prediction modes and then selects the prediction mode that produces the optimum result.”**

159. In my opinion, *Hanami* discloses the elements of claim 39.

160. For example, as discussed above with respect to claim elements 1(b) and 1(c), *Hanami* discloses a motion estimation search system that performs motion estimation in parallel (“concurrently”) using a frame predictive mode, an even field predictive mode, and an odd field predictive mode (“a plurality of different motion estimation prediction modes”) and then selects the prediction mode that produces the optimum result (e.g., based on which evaluation value corresponding to the different predictive modes is the minimum value). (*See above* Sections IX.A.1(b)-(c).)

8. Claim 41

a) **“A method, including the steps of:”**

161. I have been asked to assume that the preamble of claim 41 is limiting. Under that assumption, it is my opinion that *Hanami* discloses this element for at least the reasons given above with respect to claim element 1(a). (*See above* Section IX.A.1(a).)

b) **“concurrently performing motion estimation using each of a plurality of different motion estimation prediction modes;”**

162. *Hanami* discloses this element for at least the reasons given above with respect to claim element 1(b). (*See above* Section IX.A.1(b).)

c) **“comparing the results produced using each different prediction mode; and,”**

163. *Hanami* discloses this element for at least the reasons given above with respect to claim element 1(b). (*See above* Section IX.A.1(b).)

d) **“selecting the prediction mode that produced an optimum result; and,”**

164. *Hanami* discloses this element for at least the reasons given above with respect to claim elements 1(b)-1(c). (*See above* Sections IX.A.1(b)-(c).)

e) **“generating one or more motion vectors using the selected prediction mode.”**

165. *Hanami* discloses this element for at least the reasons given above with respect to claim element 1(d). (*See above* Section IX.A.1(d).)

B. *Hanami* and *Mitchell* Disclose Claims 2 and 5

166. In my opinion, *Hanami* and *Mitchell* disclose or suggest all of the features of claims 2 and 5. Below, I address each limitation of these claims.

1. Claim 2

167. As I discussed above, *Hanami* discloses that the first and second pixel arrays (e.g., *Hanami*’s template block TB1 and candidate reference blocks such as

macroblocks MB1-MB3 included in the search area) each have a size and structure defined by an MPEG standard, as required by claim 2. (*See above* Section IX.A.2.) However, I have been asked to assume that *Hanami* does not necessarily disclose this feature. Under that assumption, in my opinion, a person of ordinary skill in the art would have been motivated to modify *Hanami* such that its disclosed video encoding techniques would be compliant with, and therefore defined by, the MPEG standard. A person of ordinary skill in the art would have been motivated to do so in light of the knowledge of a person of ordinary skill in the art at the time of the alleged invention and the teachings of *Mitchell* (Ex. 1006).

168. For instance, the authoritative *Mitchell* textbook includes, as part of its Introduction chapter, a section entitled “Why standardize video compression?” (*Id.* at 38-39.) In it, *Mitchell* explains that “the MPEG standardization effort was initiated in 1988” as a result of the “several different industries [] converging on digital video technology.” (*Id.* at 38.) These industries included “[t]he computer industry [] looking to expand beyond its traditional text and graphic capabilities into interactive multimedia,” “[t]he consumer electronics industry [viewing] digital video as a means for improving the capabilities and interactivity of video games and entertainment media,” and “[t]he telecommunications industry [] taking advantage of the maturing compression technology to standardize teleconferencing.” (*Id.*) As

Mitchell explains, “[w]ith the sharing of technology among these diverse industrial interests, ***standardization was a natural development***” as “the cost reduction implicit in having a common digital video technology ***for all*** has been a key driving force for standardization.” (*Id.* (emphasis added).) *Mitchell* also explains that the interests of different industries converging which resulted into a common standard “lower[ed] the barriers that are usually present to the deployment of new technology.” (*Id.*)

169. In fact, *Mitchell*’s description of the need and desire to standardize video coding via MPEG is confirmed by the ’005 patent itself which states that “[t]he ***most widely accepted international standards have been proposed by the Moving Pictures Expert Group (MPEG)***, and are generally referred to as the ***MPEG-1 and MPEG-2 standards***.” (Ex. 1001 at 1:20-23 (emphasis added); Ex. 1006 at 38-39 (*Mitchell* describing standardization with respect to MPEG-1 and MPEG-2).) Other contemporaneous documents such as *Lam* (Ex. 1007) also confirm that the MPEG standard was a “commonly used video compression standard” at the time of the alleged invention. (Ex. 1007 at 1:37-40.) As one IEEE article noted, the development of MPEG-1 and MPEG-2 “resulted in an explosion of products and

services that created consumer video technology as we know it today.” (Ex. 1011 at 3.)²⁷

170. In view of the foregoing, a person of ordinary skill in the art would have been motivated to modify *Hanami* such that its video coding techniques would be compliant with, and therefore defined by, the MPEG standard. Such a person of ordinary skill in the art would have been motivated to modify *Hanami* in this way for several reasons.

171. For instance, as I discussed above, there was a need and desire to standardize video coding techniques so the different industries converging on digital video technology could share their technologies. (Ex. 1006 at 38.) In fact, there was an apparent benefit for such standardization because as *Mitchell* explains, “[i]f standardization of digital video could be achieved, the potentially large sales volumes in the consumer market would lower the cost of VLSI chips for all of the industries” and this “cost reduction” was “a key driving force for standardization.” (*Id.*) And a person of ordinary skill in the art would have also recognized that employing the standardized MPEG technology in *Hanami*’s video coding techniques would have “lower[ed] the barriers that are usually present to the deployment of new

²⁷ For Exhibit 1011, I cite to the exhibit page numbers appended to the exhibit.

technology.” (*Id.*) Moreover, as a person of ordinary skill in the art would have known at the time of the alleged invention (and as acknowledged by the ’005 patent itself), MPEG was the most widely acceptable video coding standardization body at the time of the alleged invention. Thus, a person of ordinary skill in the art would have been motivated to ensure that any video coding techniques (including those disclosed by *Hanami*) would be compliant with, and therefore defined by, MPEG. In so doing, *Hanami*’s template block TB1 and reference candidate blocks (e.g., MB1-MB3) included in the search area would have a structure and size defined by an MPEG standard.

172. In my opinion, a person of ordinary skill in the art would have had the capability and knowledge to implement the above modification to *Hanami* while ensuring that *Hanami*’s video coding techniques performed properly and would have expected the combined *Hanami-Mitchell* process to perform efficient and effective video coding without detracting from the processes disclosed by *Hanami*. This is confirmed by the fact that, as I discussed above, the data structure and block size of *Hanami*’s image signals are identical to the data structure and block size defined by MPEG. (*See above* Section IX.A.2.) Thus, in my opinion, implementing the above-discussed modification would have been a predictable application of known techniques (e.g., the video coding techniques disclosed by *Hanami* and the

standardization of such techniques via MPEG as disclosed by *Mitchell*) to achieve the foreseeable result of providing video coding techniques standardized and defined by MPEG such that the techniques would be widely applicable.

2. Claim 5

173. As I discussed above, *Hanami* discloses that that the motion coding in *Hanami* is in accordance with an MPEG standard, as required by claim 5. (*See above* Section IX.A.4.)

174. However, I have been asked to assume that *Hanami* does not necessarily disclose this feature. Under that assumption, in my opinion, a person of ordinary skill in the art would have been motivated to modify *Hanami* such that its disclosed video encoding techniques would be compliant with, and therefore defined by, the MPEG standard. In other words, for all the reasons discussed above with respect to claim 2, a person of ordinary skill in the art would have been motivated to modify *Hanami* such that its motion coding would have been in accordance with the MPEG standard in light of the knowledge of a person of ordinary skill in the art at the time of the alleged invention and the teachings of *Mitchell* (Ex. 1006). (*See above* Section IX.B.1.)

175. This is further confirmed by *Hanami* for two additional reasons. First, as discussed above with respect to claim 5, the variable length coding list for motion

vectors in *Hanami* and the MPEG-2 Part standard is identical. (*See above* Section IX.A.4.) Second, as discussed below, *Hanami*'s video encoder already includes all the functionality necessary to implement motion coding in accordance with an MPEG standard. For instance, as I discussed above with respect to claim 3, for example, *Hanami* discloses an "overall structure of an encoder for coding image data in accordance with a conventional motion-compensated predictive coding method."²⁸ (Ex. 1005 at 1:44-47.) As shown below in Figure 83, the encoder includes a preprocessing circuit 910, ***a source coding circuit 912 for executing elimination of redundancy*** and quantization on the signal preprocessed by the preprocessing circuit 910, a video multiplex coding circuit 914 for coding the signal from the source coding circuit 912, a transmission buffer 916 for buffering the image data from the video multiplex coding circuit 914, and a transmission coding circuit 918 for adapting the image data from the transmission buffer 916 to a transmission channel. (*Id.* at 1:47-55, 2:5-9, FIG. 83.)

²⁸ *See also above* Section IX.A.1(a); *see also above* n.22.

FIG. 83 PRIOR ART

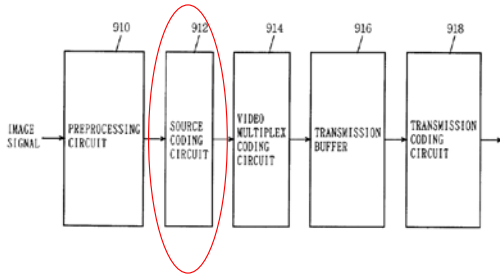
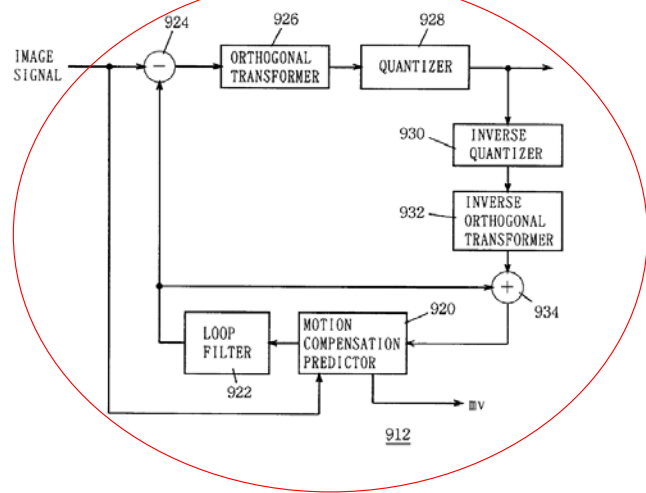


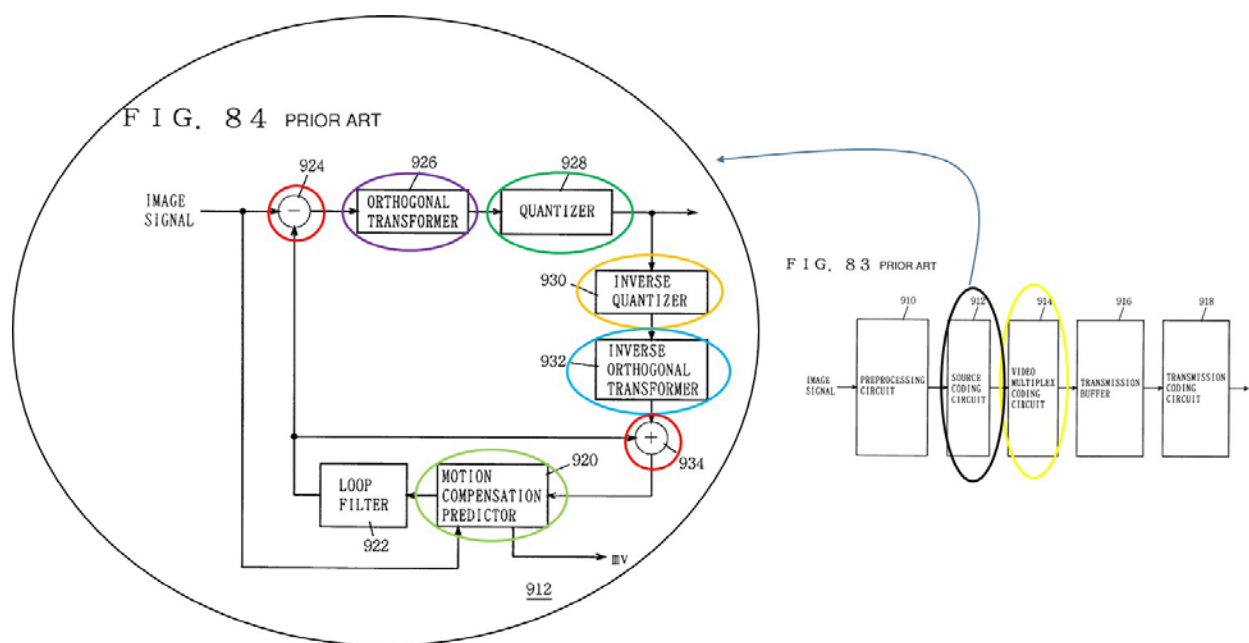
FIG. 84 PRIOR ART



(*Id.* at FIG. 83 (left) (annotated), FIG. 84 (right) (annotated).)

176. The source coding circuit 912 and the video multiplex coding circuit 914 are collectively responsible for the motion coding in *Hanami*. (*Id.* at 1:49-52; *see also id.* at 1:61-65 (disclosing that the source coding circuit 912 performs motion compensation on the input signal), 2:13-64 (disclosing with reference to Figure 84 (annotated above) that the source coding circuit includes a motion compensation predictor 920 for detection of motion vectors with respect to the input image signal), 1:66-2:4 (disclosing with reference to Figure 83 that the video multiplex coding circuit 914 applies variable length coding to the output received from the source coding circuit 912 and multiplexes this coded output).)

177. *Hanami* shows by way of its Figures 83-84 that its source coding circuit 912 and video multiplex coding circuit 914 include all the functionality necessary to implement motion coding in accordance with an MPEG standard.



(*Id.* at FIG. 84 (left) (annotated), FIG. 83 (right) (annotated).)

178. For example, with reference to an annotated version of *Hanami*'s Figure 84 reproduced above, *Hanami* discloses that its source coding circuit 912 includes:

- a motion compensation predictor 920 [annotated in light green], which includes a frame memory, for detecting motion vectors with respect to the input image signal supplied from the preprocessing circuit 910 and generating

a reference image which is motion-compensated in accordance with the motion vectors (Ex. 1005 at 2:14-18, 2:28-33),

- a loop filter 922 for filtering reference image pixel data from the motion compensation predictor 920 (*id.* at 2:19-20, 2:33-34),
- a subtracter 924 [annotated in red] for obtaining the difference between an image signal outputted from the loop filter 922 and the input image signal (*id.* at 2:20-22, 2:40-43),
- an orthogonal transformer 926 [annotated in purple] for orthogonally transforming an output of the subtracter 924 (*id.* at 2:22-24, 2:43-44),
- a quantizer 928 [annotated in green] for quantizing the data orthogonally transformed by the orthogonal transformer 926 (*id.* at 2:24-27),
- an inverse quantizer 930 [annotated in orange] for transforming the data quantized in the quantizer 928 to the signal state before the quantization (*id.* at 2:46-48, 2:52-55),
- an inverse orthogonal transformer 932 [annotated in blue] for performing inverse orthogonal transformation on an output of the inverse quantizer 930 (*id.* at 2:48-55), and
- an adder 934 [annotated in red] for adding up outputs of the loop filter 922 and the inverse orthogonal transformer 932 (*id.* at 2:51-52, 2:58-61).

179. Similarly, with reference to an annotated version of *Hanami*'s Figure 83 reproduced above, *Hanami* discloses that its video multiplex coding circuit 914 [annotated in yellow], which receives the output of the source coding circuit 912, converts the output image data from the quantizer 928 and the motion vectors from the motion compensation predictor 920 to variable-length codes and thereafter multiplexes them (*id.* at 1:66-2:4, 2:61-64).

180. In my opinion, a person of ordinary skill in the art would have understood based on the above configuration of *Hanami*'s source coding circuit 912 and video multiplex coding circuit 914 that *Hanami*'s circuitry includes all the required components to perform motion coding in accordance with an MPEG standard, as recited in claim 5. For example, a person of ordinary skill in the art at the time of the alleged invention would have known that to perform motion encoding in accordance with an MPEG standard, the encoder would need to include components such as *Hanami*'s motion compensation predictor 920, subtracter 924, orthogonal transformer 926, quantizer 928, inverse quantizer 930, inverse orthogonal transformer 932, adder 934, and video multiplex coding circuit 914.

181. My understanding is supported by contemporaneous MPEG-related documentation at the time of the alleged invention such as U.S. Patent No. 6,035,069

to Nagai *et al.* (“Nagai”) (Ex. 1013).²⁹ For instance, with reference to Figure 1, Nagai discloses an example of a conventional moving-picture coding system using the motion-compensated adaptive prediction and discrete cosine transform implemented by international standardized methods such as MPEG-1 and MPEG-2.

In an image transmitting and/or storing system such as a video telephone, a videoconferencing system, a portable information terminal, a digital video disc system and a television broadcasting system, various systems, which include the motion compensation, the discrete cosine transform, the sub-band coding, the pyramid coding and the combinations thereof, have been developed as

²⁹ In this portion of my analysis, I rely solely on the teachings of *Hanami* and *Mitchell* for disclosing the features of claim 2. I discuss *Nagai* (Ex. 1013) to demonstrate the knowledge of a person of ordinary skill in the art at the time of the alleged invention for the '005 patent regarding the conventional configuration and layout of an MPEG encoder. Given such knowledge, such a skilled person, in my opinion, would have understood that *Hanami* already includes all the required components for an MPEG encoder to perform motion coding and *Mitchell* supplies the motivation to modify *Hanami*'s encoder such that it would perform motion coding in accordance with an MPEG standard.

techniques for compressing and coding an image data into a smaller amount of data for transmission/storing. *In addition, as international standard methods for compressing and coding a moving picture, the methods "ISO•MPEG1", "ISO•MPEG2", "ITU-T•H. 261", and "TU-T•H.262" are standardized.* All these methods are compressing and coding methods, *each comprising a combination of the motion-compensated adaptive prediction and the discrete cosine transform ...*

As an example of conventional moving-picture coding systems, the basic construction of a coding system using the motion-compensated adaptive prediction and the discrete cosine transform is shown in FIG. 1.

(Ex. 1013 at 1:21-43 (emphasis added).)

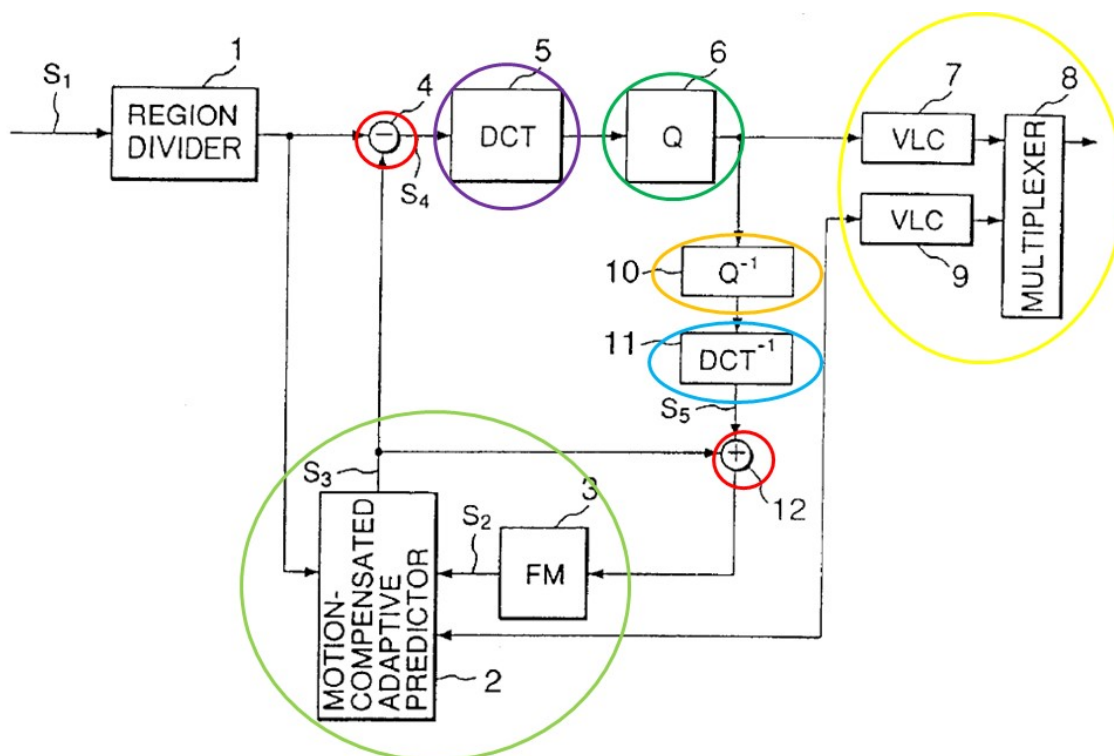


FIG.1
PRIOR ART

(*Id.* at FIG. 1 (annotated).)

182. As shown above, like *Hanami*'s overall encoder configuration, a conventional MPEG-1/MPEG-2 moving-picture coding system would have included:

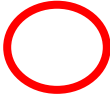
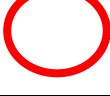
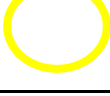
- a motion-compensated adaptive predictor 2 [annotated in light green], like *Hanami*'s motion compensation predictor 920, which detects a motion vector between the input image signal S1 and a reference image signal S2 of the previous frame (stored in a frame memory 3 and which has been already coded

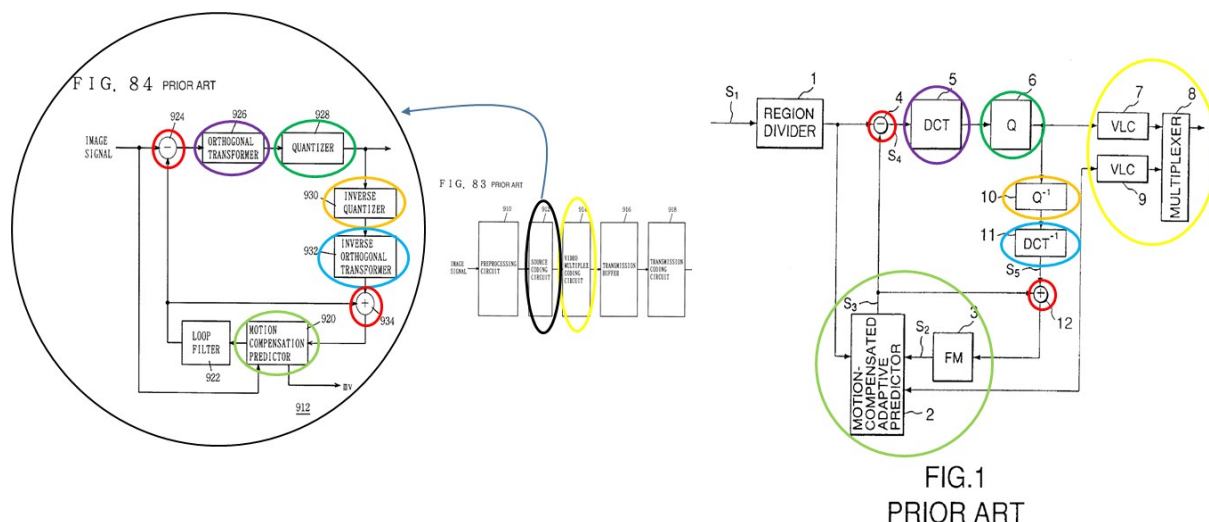
and locally decoded), and performs the motion compensation with respect to the reference image signal using the detected motion vector, so as to produce a prediction signal (*id.* at 1:46-53),

- a subtracter 4 [annotated in red], like *Hanami*'s subtracter 924, which subtracts the prediction signal S3 from the input image signal S1 to output a predictive residual signal S4 (*id.* at 1:60-62),
- a discrete cosine transformer 5 [annotated in purple], like *Hanami*'s orthogonal transformer 926, which carries out the discrete cosine transform (DCT) of the predictive residual signal S4 (*id.* at 1:62-65),
- a quantizer 6 [annotated in green], like *Hanami*'s quantizer 928, which quantizes the DCT coefficient obtained by the discrete cosine transform (*id.* at 1:65-67),
- an inverse quantizer 10 [annotated in orange], like *Hanami*'s inverse quantizer 930, which inverse-quantizes the quantized DCT coefficients output by the quantizer 6 (*id.* at 2:5-9),
- an inverse discrete-cosine transformer 11 [annotated in blue], like *Hanami*'s inverse orthogonal transformer 932, for performing inverse discrete-cosine transform on an output of the inverse quantizer 10 (*id.* at 2:5-9),

- an adder 12 [annotated in red], like *Hanami*'s adder 924, for adding up the output of the inverse discrete-cosine transformer 11 to the adaptive prediction signal S3 to generate a locally decoded signal to be stored in the frame memory 3 (*id.* at 2:9-12), and
- variable-length encoders 7, 9, and multiplexer 8 [annotated in yellow], like *Hanami*'s video multiplex coding circuit 914, where the variable-length encoder 7 variable-length encodes the quantized DCT coefficients output from quantizer 6, the variable-length encoder 9 variable-length encodes the motion vector output from the motion-compensated adaptive predictor 2, and the multiplexer 8 which multiplexes the output from the encoders 7, 9 (*id.* at 2:1-5).

The table below shows the correspondence between these components:

Annotation color	<i>Nagai</i>	<i>Hanami</i>
	motion- compensated adaptive predictor 2	motion compensation predictor 920
	subtractor 4	subtractor 924
	discrete cosine transformer 5	orthogonal transformer 926
	quantizer 6	quantizer 928
	inverse quantizer 10	inverse quantizer 930
	inverse discrete- cosine transformer 11	inverse orthogonal transformer 932
	adder 12	adder 924
	variable-length encoders 7, 9, and multiplexer 8	video multiplex coding circuit 914



(Ex. 1005 at FIG. 84 (left) (annotated), FIG. 83 (center) (annotated); Ex. 1013 at FIG. 1 (right) (annotated).)

183. As can be seen above based on a side-by-side comparison of the motion coding components present in *Hanami* and the motion coding components in a conventional MPEG decoder, *Hanami* includes all the required components in its encoder to perform motion coding in accordance with an MPEG standard. In particular, aside from the inclusion of a loop filter 22 in *Hanami*'s encoder shown in Figure 84, all the other blocks/circuitry shown in Figure 84 would have been compliant with an MPEG standard such as MPEG-1 or MPEG-2. (Ex. 1013 at 21-43.) To the extent that the loop filter 22 would not have been part of an MPEG standard, a person of ordinary skill in the art would have been motivated to modify *Hanami* to exclude the loop filter 22 since it is provided to merely filter reference image pixel data and this would be a matter of design choice. For instance, while

the loop filter 22 may provide some benefit, a person of ordinary skill in the art would have understood that the trade-off between having a loop filter as opposed to making the overall encoder compliant with the MPEG standard (by excluding the loop filter 22) would have weighed in favor of excluding the loop filter 22 for all the reasons discussed above with respect to claim 2 relating to the benefits of making a video coding technique compliant with the MPEG standard. (*See above* Section IX.B.1.)

184. For example, by ensuring that *Hanami*'s encoder performs motion coding in accordance with an MPEG standard, a person of ordinary skill in the art would have addressed the known need and desire at the time of the alleged invention to standardize video coding techniques so the different industries converging on digital video technology could share their technologies. (Ex. 1006 at 38.) In fact, there was an apparent benefit for such standardization because as *Mitchell* explains, "[i]f standardization of digital video could be achieved, the potentially large sales volumes in the consumer market would lower the cost of VLSI chips for all of the industries" and this "cost reduction" was "a key driving force for standardization." (*Id.*) And a person of ordinary skill in the art would have also recognized that employing the standardized MPEG technology in *Hanami*'s video coding techniques would have "lower[ed] the barriers that are usually present to the deployment of new

technology.” (*Id.*) Moreover, as a person of ordinary skill in the art would have known at the time of the alleged invention (and as acknowledged by the ’005 patent itself), MPEG was the most widely acceptable video coding standardization body at the time of the alleged invention. Thus, a person of ordinary skill in the art would have been motivated to ensure that any video coding techniques (including those disclosed by *Hanami*) would be compliant with, and therefore defined by, MPEG. In so doing, *Hanami*’s motion coding performed by the blocks/circuitry discussed above with respect to Figures 83-84 would have performed in accordance with an MPEG standard.

185. In my opinion, a person of ordinary skill in the art would have had the capability and knowledge to implement the above modification to *Hanami* while ensuring that *Hanami*’s video coding techniques performed properly and would have expected the combined *Hanami-Mitchell* process to perform efficient and effective video coding without detracting from the processes disclosed by *Hanami*. This is confirmed by the fact that, as I discussed above, *Hanami* already includes the requisite circuitry to perform motion coding in accordance with an MPEG standard. (*See above* Section IX.A.2.) Thus, in my opinion, implementing the above-discussed modification would have been a predictable application of known techniques (e.g., the video coding techniques disclosed by *Hanami* and the well-

known motion coding technique defined by an MPEG standard) to achieve the foreseeable result of providing video coding techniques (including motion coding) to be performed in accordance with an MPEG standard. (*Id.*)

X. CONCLUSION

186. I declare that all statements made herein of my knowledge are true, and that all statements made on information and belief are believed to be true, and 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.

Dated: September 9, 2019

By: Majid Rabbani
Majid Rabbani, Ph.D.