

**UNITED STATES INTERNATIONAL TRADE COMMISSION
Washington, DC 20436**

**Before the Honorable MaryJoan McNamara
Administrative Law Judge**

In the Matter of

**CERTAIN ELECTRONIC DEVICES,
INCLUDING SMARTPHONES,
COMPUTERS, TABLET COMPUTERS,
AND COMPONENTS THEREOF**

Inv. No. 337-TA-1373

RESPONDENTS' OPENING *MARKMAN* BRIEF

TABLE OF CONTENTS

I.	INTRODUCTION	1
II.	APPLICABLE LEGAL PRINCIPLES.....	3
A.	Claim Construction	3
B.	Means-Plus Function Limitations	5
C.	Whether the Preamble is Limiting	7
III.	OVERVIEW OF THE ASSERTED PATENTS.....	8
A.	'877 Patent	8
1.	Overview	8
2.	The Asserted Claims	10
B.	'859 Patent	12
1.	Entropy Encoding Background.....	12
2.	Admitted Entropy Coding Prior Art	13
3.	Alleged Invention of the '859 Patent	15
4.	The Asserted Claims of the '859 Patent	20
C.	'556 Patent	21
1.	Background	21
2.	Sparse De-noising Filters	25
3.	Alleged Invention of the '556 Patent	28
4.	The Asserted Claims of the '556 Patent	29
IV.	LEVEL OF ORDINARY SKILL IN THE ART	30
V.	'877 PATENT	30
A.	“circuit configured to” (claims 7, 8)	30
VI.	'859 PATENT	33

A.	“arithmetic decoding method for symbols coded in the form of a stream, comprising the following steps applied at switching points distributed in said stream to decode a current symbol” (claim 10).....	33
1.	The Preamble of Claim 10 is the Language Before “Comprising”	34
2.	The Preamble of Claim 10 is Limiting	35
3.	The Construction of “Switching Points”.....	39
B.	“means for” terms of claim 15	47
1.	“means for decoding from the stream, at switching points distributed in said stream to decode a current symbol, for said current symbol, a probability model identifier” (claim 15)	47
2.	“means for selecting using said probability model identifier a probability model in a set comprising a current probability model defined from symbols coded previous to the current symbol and a subset of probability models wherein subset comprises at least one probability model defined by default” (claim 15).....	56
3.	“means for adding in said subset said current probability model, the selected probability model becoming the current probability model” (claim 15).....	60
4.	“means for decoding said current symbol with said current probability model” (claim 15).....	63
5.	“means for updating said current probability model according to the decoding of said current symbol” (claim 15).....	65
VII.	’556 PATENT	68
A.	“sparse de-noising filter”	68
1.	InterDigital Disclaimed “Deringing Filters” During Prosecution	69
2.	Alternatively, Sparse De-noising Filter Should Be Construed According to the Specification.....	76
B.	“adaptive sparse de-noising filter”.....	80
VIII.	CONCLUSION.....	82

TABLE OF AUTHORITIES

	Page(s)
Cases	
<i>Advanced Tech. Incubator, Inc. v. Sharp Corp.</i> , 2009 WL 749911 (E.D. Tex. 2009)	41
<i>Aristocrat Techs. Austl. Pty Ltd. V. Int’l Game Tech.</i> , 521 F.3d 1328, 1333 (Fed. Cir. 2008).....	35
<i>Aspex Eyewear, Inc. v. Marchon Eyewear, Inc.</i> , 672 F.3d 1335 (Fed. Cir. 2012).....	33
<i>Bennett Marine, Inc. v. Lenco Marine, Inc.</i> , 549 Fed. Appx. 947 (Fed. Cir. 2013).....	50, 59, 65, 68
<i>Bicon, Inc. v. Straumann Co.</i> , 441 F.3d 945 (Fed. Cir. 2006).....	7, 35
<i>Bio-Rad Labs., Inc. v. 10X Genomics Inc.</i> , 967 F.3d 1353 (Fed. Cir. 2020).....	37
<i>Bondyopadhyay v. United States</i> , 748 F. App’x 301 (Fed. Cir. 2018)	7
<i>Catalina Mktg. Int’l, Inc. v. Coolsavings.com, Inc.</i> , 289 F.3d 801 (Fed. Cir. 2002).....	<i>passim</i>
<i>Certain Computer Network Security Equip. and Systems</i> , Inv. No. 337-TA-1314, Order No. 37 (USITC Feb. 22, 2023) (McNamara, ALJ)	33
<i>Certain LTE- & 3G-Compliant Cellular Commc’ns Devices</i> , Inv. No. 337-TA-1138, Order No. 50, USITC Pub. No. 5258 (USITC Jul. 15, 2019).....	4
<i>CIAS, Inc. v. Alliance Gaming Corp.</i> , 504 F.3d 1356 (Fed. Cir. 2007).....	35
<i>Computer Docking Station Corp. v. Dell, Inc.</i> , 519 F.3d 1366 (Fed. Cir. 2008).....	75
<i>Corning Glass Works v. Sumitomo Elec. U.S.A.</i> , 868 F.2d at 1257	38, 39
<i>Default Proof Credit Card Sys., Inc. v. Home Depot U.S.A., Inc.</i> , 412 F.3d 1291 (Fed. Cir. 2005).....	6

<i>EMA Electromechanics, Inc. v. Siemens Corp.</i> , No. 6:21-CV-1001-ADA, 2022 WL 526244 (W.D. Tex. Feb. 16, 2022).....	50, 52, 54, 55
<i>Ergo Licensing, LLC v. CareFusion 303, Inc.</i> , 673 F.3d 1361 (Fed. Cir. 2012).....	53, 55, 56
<i>Finisar Corp. v. DirecTV Group, Inc.</i> , 523 F.3d 1323 (Fed. Cir. 2008).....	56, 57
<i>In re Fought</i> , 941 F.3d 1175 (Fed. Cir. 2019).....	35
<i>Function Media, L.L.C. v. Google, Inc.</i> , 708 F.3d 1310 (Fed. Cir. 2013).....	<i>passim</i>
<i>Gillette Co. v. Energizer Holdings</i> , 405 F.3d 1367, 1370 (Fed. Cir. 2005).....	4, 35
<i>Grober v. Mako Prods.</i> , 686 F.3d 1335 (Fed. Cir. 2012).....	75
<i>Lockheed Martin Corp. v. Space Sys./Loral, Inc.</i> , 324 F.3d 1308 (Fed. Cir. 2003).....	50
<i>Luminara Worldwide, LLC v. Liown Elecs. Co. Ltd.</i> , 814 F.3d 1343 (Fed. Cir. 2016).....	44
<i>Markman v. Westview Instruments, Inc.</i> , 52 F.3d 967 (Fed. Cir. 1995), <i>aff'd</i> , 517 U.S. 370 (1996)	1, 3, 5
<i>Morpho Komodo LLC v. Blu Prods. Inc.</i> , 2016 WL 3976491 (E.D. Tex. 2016)	41
<i>Noah Sys., Inc. v. Intuit Inc.</i> , 675 F.3d 1302 (Fed. Cir. 2012).....	<i>passim</i>
<i>Omega Eng'g, Inc. v. Raytek Corp.</i> , 334 F.3d 1314 (Fed. Cir. 2003).....	71, 75, 78
<i>Phillips v. AWH Corp.</i> , 415 F.3d 1303 (Fed. Cir. 2005) (en banc).....	4, 5
<i>Princeton Digital Image Corp. v. Konami Digital Entm't Inc.</i> , 2017 WL 6375173 (D. Del. 2017)	<i>passim</i>
<i>Proveris Sci. Corp. v. Innovasystems, Inc.</i> , 739 F.3d 1367 (Fed. Cir. 2014).....	7

<i>Radware Ltd. v. A10 Networks, Inc.</i> , No. C-13-02024-RMW, 2014 WL 1572644 (N.D. Cal. Apr. 18, 2014).....	33
<i>Rowe v. Dror</i> , 112 F.3d 473 (Fed. Cir. 1997).....	7
<i>In re Smith Int'l, Inc.</i> , 871 F.3d 1375 (Fed. Cir. 2017).....	4
<i>Starhome GmbH v. AT&T Mobility LLC</i> , 743 F.3d 849 (Fed. Cir. 2014).....	39, 46, 48
<i>Synchronoss Techs., Inc. v. Dropbox, Inc.</i> , 987 F.3d 1358 (Fed. Cir. 2021).....	55, 56
<i>Tanabe Seiyaku Co., Ltd. V. U.S. Int'l Trade Comm'n</i> , 109 F.3d 726 (Fed. Cir. 1997).....	39, 46
<i>Third Wave Techs., Inc. v. Eragen Bioscience, Inc.</i> , 2003 WL 23100277 (W.D. Wis. 2003).....	41
<i>Trading Techs. Int'l v. Open E Cry</i> , LLC, 728 F.3d 1309 (Fed. Cir. 2013).....	71
<i>Verve v. Crane Cams</i> , 311 F.3d 1116 (Fed. Cir. 2002).....	5
<i>Vitrionics Corp. v. Conceptronic, Inc.</i> , 90 F.3d 1576 (Fed. Cir. 1996).....	5
<i>Williamson v. Citrix Online, LLC</i> , 792 F.3d 1339 (Fed. Cir. 2015).....	5, 6, 53, 58
Statutes	
35 U.S.C. § 112 ¶ 6.....	<i>passim</i>

TABLE OF EXHIBITS

Exhibit No.	Description
JXM-0001	U.S. Patent No. 10,250,877
JXM-0002	U.S. Patent No. 8,674,859
JXM-0003	U.S. Patent No. 9,674,556
RXM-0001	U.S. Patent 7,932,843 B2
RXM-0002	U.S. Patent 8,351,502 B2
RXM-0003	U.S. Patent 8,416,857 B2
RXM-0004	U.S. Patent 8,718,149 B2
RXM-0005	U.S. Patent 9,008,171 B2
RXM-0006	U.S. Patent ,9386,316 B2
RXM-0007	File History for EP2449683
RXM-0021	File History for U.S. Pat. App. No. 15/585462
RXM-0022	File History for U.S. Pat. No. 9,277,243
RXM-0023	File History for European Pat. Pub. No. 2 082 583
RXM-0025	Dabov, K., 2007, August. Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering. IEEE Transactions on Image Processing (Vol. 16, pp 2080-95). IEEE.
RXM-0026	Excerpts to Richardson, I., 2010. The H.264 Advanced Video Compression Standard (2 nd Ed.), WILEY.
RXM-0033	Ravishankar, S., 2015, June. Online Sparsifying Transform Learning—Part I: Algorithms. IEEE Journal of Selected Topics in Signal Processing (Vol. 9, No. 4, pp 625-636). IEEE.
RXM-0034	Ravishankar, S., 2013, March. Learning Sparsifying Transforms. IEEE Transactions on Image Processing (Vol. 61, No. 5, pp 1072-86). IEEE.
RXM-0035	U.S. Pat. App. No. 2011/0222597
RXM-0038	Dabov, K. 2007, Sept. Video Denoising by Sparse 3D Transform-Domain Collaborative Filtering. 15th European Signal Processing Conference (EUSIPCO 2007) 15th European Signal Processing Conference (EUSIPCO 2007). (pp 145–49). EURASIP.
RXM-0040	Xu, J., 2009. Sparsity-Based Deartifacting Filtering in Video Compression. ICIP 2009 (pp 3933–36). IEEE.
RXM-0041	Declaration of Michael Orchard in support of Respondents’ Opening Markman Brief
RXM-0042	US Patent Appl. No. 2003/0219073A1
RXM-0043	US Patent 9,277,243 B2
RXM-0044	Ex. 5 to InterDigital’s First Suppl. Responses to Respondents’ First Set of Interrogatories

I. INTRODUCTION

Respondents Lenovo PC HK Limited, Lenovo (United States) Inc., and Motorola Mobility LLC (collectively, “Respondents”) respectfully submit this opening *Markman* brief in support of their proposed constructions for the disputed claim terms within U.S. Patent Nos. 10,250,877 (“’877 patent”), 8,674,859 (“’859 patent”), and 9,674,556 (“’556 patent”). Respondents’ proposed constructions reflect how a person of ordinary skill in the art (“POSITA”) as of the relevant timeframe would have understood the terms in question in light of the intrinsic and extrinsic evidence, as well as representations the applicants themselves have made to U.S. and foreign patent agencies.

For the ’877 patent, the only dispute is the plain and ordinary meaning of “circuit configured to.” At one point, the parties had all agreed that it means “requires pre-existing programming of hardware and software to perform the cited functionality,” which is almost identical to what the ALJ adopted for the same term in Inv. No. 1314. That was the construction Complainants had identified in the Joint Claim Construction Chart, and the parties all agreed to it during the subsequent meet and confer process. Later, Complainants changed their mind and proposed a rearranged version of the language that moves “pre-existing” to a different part of the phrase, to modify “hardware and software” instead of “programming” (i.e., “requires *pre-existing* hardware and software programmed to perform the cited functionality”). But Complainants have not explained why they made this change, how it supposedly different from the agreement the parties had already reached, or why it is justified here.

For the ’859 patent, the parties dispute which portion of claim 10 of the ’859 patent is the preamble and how the ALJ should construe “switching points.” Respondents and Staff rely on well-settled Federal Circuit law providing that claim language after “transition” terms like “comprising” is part of the claim body. Complainants have not explained why they believe that

law does not apply here. With respect to “switching points,” the parties all agree the specification includes a passage that describes what “switching points” are. But because the parties disagree on what that passage means, they have a dispute over claim scope that needs to be resolved during claim construction. Respondents’ construction clarifies that “switching points” are located at points in the stream where “a probability model can be selected,” which is the entire point of the alleged invention and is the improvement the ’859 patent claims to make over the prior art. Indeed, Complainants *themselves* have argued for that interpretation of “switching points” outside the context of this litigation, in counterpart proceedings before the European Patent Office. Here, on the other hand, Complainants apply an overly-broad interpretation of “switching points” that is divorced from the intrinsic record and inconsistent with positions they have previously taken.

The parties also dispute certain “means for” terms in the ’859 patent. There is no dispute these terms are written in “means-plus-function” format and are subject to 35 U.S.C. § 112 ¶ 6. The main dispute is whether the ’859 patent provides a sufficient description of corresponding structure, which is required to render the terms definite. It does not. The ’859 patent describes the claimed functions almost entirely in *functional* terms. What little information it provides about where the claimed functions would be performed, such as a “decoding device” or “entropy decoding module,” is entirely generic. As abundant extrinsic evidence confirms, a POSITA would have understood that such decoding devices or modules can be implemented in any number of different structures, such as a microprocessor, discrete logic, an application specific integrated circuit, a field programmable gate array, or digital signal processor, to name a few. The specification does not describe which structures would be used, however, rendering the terms indefinite under Federal Circuit law. This deficiency cannot be remedied by testimony from Complainants’ expert, because “the testimony of one of ordinary skill in the art cannot supplant

the total absence of structure from the specification.” *Noah Sys., Inc. v. Intuit Inc.*, 675 F.3d 1302, 1312 (Fed. Cir. 2012) (quotation omitted). It is the *specification*, not expert gloss, that must disclose corresponding structure, *id.*, and here it does not.

Finally, the parties dispute the meaning of “sparse de-noising filter” in the ’556 patent’s claims. Respondents’ construction reflects the clear disclaimer the applicant made during prosecution of related applications, where it repeatedly distinguished prior art by representing that the claimed “sparse de-noising filter” is *not a deringing filter*. Complainants are bound by those representations here. Without any support from the intrinsic evidence, Complainants propose “a filter that reduces noise based on a sparse representation of the signal,” a construction that adds more ambiguity than it resolves. If further construction is necessary beyond the prosecutorial disclaimer, Respondents propose an alternative construction based on how the inventors defined the term in the specification. Finally, Complainants’ proposal to limit the meaning of “adaptive” in the larger term “adaptive sparse de-noising filter” restricts the claims solely to filters relying on “image region” as an input. That proposal is inconsistent with the specification, however, which indicates the term encompasses a broader range of inputs than just an “image region.”

For the reasons discussed in greater detail below, Respondents respectfully request that the ALJ adopt their proposed construction of each term.

II. APPLICABLE LEGAL PRINCIPLES

A. Claim Construction

Claim construction is a matter of law. *Markman v. Westview Instruments, Inc.*, 52 F.3d 967, 979 (Fed. Cir. 1995), *aff’d*, 517 U.S. 370 (1996). “Claims should be given their ordinary and customary meaning as understood by a person of ordinary skill in the art, viewing the claim terms in the context of the entire patent.” *Certain LTE- & 3G-Compliant Cellular Commc’ns Devices*,

Inv. No. 337-TA-1138, Order No. 50, USITC Pub. No. 5258 at *121 (USITC Jul. 15, 2019) (citing *Phillips v. AWH Corp.*, 415 F.3d 1303, 1312–13 (Fed. Cir. 2005) (en banc)).

“Claim construction requires [a] court to place the claim language in its proper technological and temporal context. The best tools for this enterprise are the various forms of intrinsic evidence and, when appropriate, extrinsic evidence.” *Gillette Co. v. Energizer Holdings, Inc.*, 405 F.3d 1367, 1370 (Fed. Cir. 2005) (citation omitted). The claims themselves provide substantial guidance with regard to the meaning of disputed claim language. *Phillips*, 415 F.3d at 1314. Beyond the claim language, the specification is “the single best guide to the meaning of a disputed term” and is usually dispositive. *Id.* at 1315. Thus, claim terms should be construed in light of the specification “with a view to ascertaining the invention.” *Id.* at 1316 (citation omitted). “The correct inquiry” for purposes of claim construction “is not whether the specification proscribes or precludes some broad reading of the claim term.” *In re Smith Int’l, Inc.*, 871 F.3d 1375, 1382–83 (Fed. Cir. 2017). Rather, the correct construction must “correspond[] with . . . how the inventor describes his invention in the specification, i.e., an interpretation that is consistent with the specification.” *Id.* at 1383 (citation and internal quotation marks omitted). The prosecution history may also explain the meaning of claim language and reveal “how the inventor understood the invention and whether the inventor limited the invention in the course of prosecution, making the claim scope narrower than it would otherwise be.” *Phillips*, 415 F.3d at 1317 (citation omitted).

In addition to intrinsic evidence, courts are also authorized to consider extrinsic evidence, including “expert and inventor testimony, dictionaries, and learned treatises.” *Id.* (citation omitted). Such extrinsic evidence is particularly appropriate where it helps the court “‘to better understand the underlying technology’ and the way in which one of skill in the art might use the

claim terms.” *Id.* at 1318 (quoting *Vitrionics Corp. v. Conceptronic, Inc.*, 90 F.3d 1576, 1584 n.6 (Fed. Cir. 1996)). Such evidence, including expert opinions, can elucidate “the true meaning of language used in the patent claims.” *Markman*, 52 F.3d at 980; *Verve v. Crane Cams*, 311 F.3d 1116, 1119 (Fed. Cir. 2002) (“While reference to intrinsic evidence is primary in interpreting claims, the criterion is the meaning of words as they would be understood by persons in the field of the invention.”). Expert testimony can be of assistance if, with respect to the disputed claim language, it identifies what the “particular meaning in the pertinent field” would be to one skilled in the art. *Phillips*, 415 F.3d at 1318.

B. Means-Plus Function Limitations

A patent may describe a claim element in “means-plus-function” format, meaning the claim describes what the particular element does (its function) rather than how it is made (its structure). *See* 35 U.S.C. § 112 ¶ 6.

“Construing a means-plus-function claim term is a two-step process.” *Williamson v. Citrix Online, LLC*, 792 F.3d 1339, 1351–52 (Fed. Cir. 2015). “The court must first identify the claimed function.” *Id.* (citing *Noah Sys.*, 675 F.3d at 1311). “Then, the court must determine what structure, if any, disclosed in the specification corresponds to the claimed function.” *Id.* “If the patentee fails to disclose adequate corresponding structure, the claim is indefinite. *Id.* (citation omitted).

“Structure disclosed in the specification qualifies as ‘corresponding structure’ if the intrinsic evidence clearly links or associates that structure to the function recited in the claim.” *Id.* (citation omitted). “Even if the specification discloses corresponding structure, the disclosure must be of ‘adequate’ corresponding structure to achieve the claimed function.” *Id.* (citation omitted). For claims governed by § 112 ¶ 6, “the specification must contain sufficient descriptive text by

which a person of skill in the field of the invention would know and understand what structure corresponds to the means limitation.” *Function Media, L.L.C. v. Google, Inc.*, 708 F.3d 1310, 1317 (Fed. Cir. 2013). “[I]f a person of ordinary skill in the art would be unable to recognize the structure in the specification and associate it with the corresponding function in the claim, a means-plus-function clause is indefinite.” *Williamson*, 792 F.3d at 1352 (citation omitted).

While “the question of whether a claim is indefinite is based on how the claim limitation would be understood by one of skill in the art, ‘the testimony of one of ordinary skill in the art cannot supplant the total absence of structure from the specification.’” *Noah Sys.*, 675 F.3d at 1312 (quoting *Default Proof Credit Card Sys., Inc. v. Home Depot U.S.A., Inc.*, 412 F.3d 1291, 1302 (Fed. Cir. 2005)). “The prohibition against using expert testimony in this manner is a direct consequence of the requirement that the specification itself adequately disclose the corresponding structure.” *Noah Sys.*, 675 F.3d at 1312 (citation omitted).

In cases involving a “special purpose computer-implemented means-plus-function limitation,” the Federal Circuit has “consistently required that the structure disclosed in the specification be more than simply a general purpose computer or microprocessor.” *Id.* (citing *Aristocrat Techs. Austl. Pty Ltd. V. Int’l Game Tech.*, 521 F.3d 1328, 1333 (Fed. Cir. 2008)). Instead, the court has “require[d] that the specification disclose an algorithm for performing the claimed function.” *Id.* (internal quotation and citation omitted). “The specification can express the algorithm in any understandable terms including as a mathematical formula, in prose, or as a flow chart, or in any other manner that provides sufficient structure.” *Id.* (internal quotation and citation omitted). “Simply disclosing software, however, without providing some detail about the means to accomplish the function[,] is not enough.” *Id.* (internal quotation and citation omitted). The algorithm requirement is necessary “[b]ecause general purpose computers can be programmed

to perform very different tasks in very different ways” and “without disclosing any corresponding structure, one of skill simply cannot perceive the bounds of the invention.” *Noah Sys.*, 675 F.3d at 1318 (internal quotation and citation omitted).

C. Whether the Preamble is Limiting

A preamble is limiting if: (1) it supplies antecedent basis for terms in a claim body; (2) there is clear reliance on the preamble during prosecution; or (3) it is “necessary to give life, meaning, and vitality” to the claims. *Catalina Mktg. Int’l, Inc. v. Coolsavings.com, Inc.*, 289 F.3d 801, 808 (Fed. Cir. 2002). “The determination of whether preamble recitations are structural limitations or mere statements of purpose or use ‘can be resolved only on review of the entirety of the patent to gain an understanding of what the inventors actually invented and intended to encompass by the claim.’ [] The inquiry involves examination of the entire patent record to determine what invention the patentee intended to define and protect.” *Rowe v. Dror*, 112 F.3d 473, 478 (Fed. Cir. 1997) (citation omitted); *see also Bondyopadhyay v. United States*, 748 F. App’x 301, 306 (Fed. Cir. 2018); *Proveris Sci. Corp. v. Innovasystems, Inc.*, 739 F.3d 1367, 1373 (Fed. Cir. 2014) (“[T]he preamble of claim 3 is the only reference in any independent claim to the inventive concept of capturing a sequence of images in order to characterize the time evolution of the spray plume. This fact alone is likely sufficient to support a conclusion that the preamble is limiting.”). A preamble is also limiting “if it recites essential structure that is important to the invention or necessary to give meaning to the claim.” *Bicon, Inc. v. Straumann Co.*, 441 F.3d 945, 952 (Fed. Cir. 2006).

III. OVERVIEW OF THE ASSERTED PATENTS

A. '877 Patent

1. Overview

The '877 patent is generally directed to video encoding and decoding. JXM-1 ('877 Patent) at Abstract. Because many of the technical details are not necessary to understand the parties' claim construction dispute, Respondents provide a high-level overview below.

Video encoding and decoding techniques generally use a “reference image” to encode or decode another image (i.e., a current image). *Id.* To do so, the encoder/decoder applies two operations to the reference image. *First*, the encoder/decoder applies “resampling” to the reference image, which is changing the size of the reference image so that it matches that of the current image. The reference image may be “upsampled” to make it larger, or “downsampled” to make it smaller.¹ *Second*, the encoder/decoder also applies “motion compensation” to the reference image, which shifts the content of the reference image to line up with the current image to account for any movement between the two.

The '877 patent concedes that it was known in the prior art to both resample and motion compensate a reference image. *Id.* at 1:45–64. According to the '877 patent, however, the prior art applied these two steps separately, one after the other. *Id.* That is, prior art encoders/decoders first resampled the reference image, storing the resampled reference image in a “Decoded Picture Buffer” (DPB), and they then applied motion compensation to the resampled reference image. *See id.* at 1:24–2:13. The '877 patent asserts that this two-step process was inefficient because the

¹ Downsampling may also be referred to as “subsampling.” *See, e.g.*, JXM-1 ('877 Patent) at 1:55–59.

intermediate, resampled reference images must be stored in the DPB, which consumes memory. *Id.* at 2:8–13.

The '877 patent purports to provide an improvement over the prior art by performing the resampling and motion compensation steps jointly, in a single operation. The patent claims that this reduces memory usage because the intermediate, resampled reference images need not be stored in a DPB if the two operations are performed at once. *Id.* at 2:17–18, 5:48–52. Figure 3, reproduced and annotated below, is illustrative.

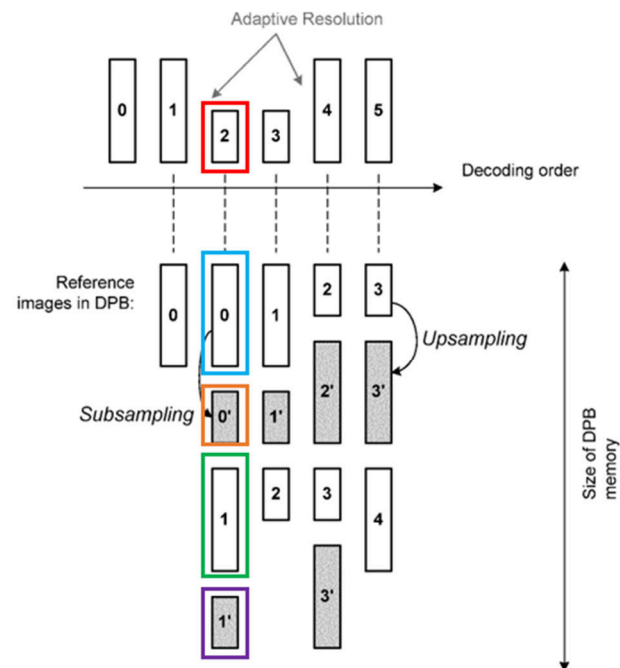


FIGURE 3

Id. at Fig. 3 (annotated). In this example from the '877 patent, image 2 (annotated above in **red**) is reconstructed using reference images 0 and 1 (annotated in **blue** and **green**, respectively). Because image 2 is smaller than references images 0 and 1, reference images 0 and 1 must be resampled prior to motion compensation. The resulting resampled images, 0' and 1' (shaded gray and annotated in **orange** and **purple**) are stored in the DPB. *See id.* at 1:24–2:13. The resampled

images 0' and 1' are then motion compensated to decode image 2. *See id.* According to the '877 patent, the purported invention makes these resampled reference images unnecessary by using a single, joint filter for resampling and motion compensation, eliminating the need to store intermediate “resampled” reference images and saving space in the DPB. *Id.* at 5:48–52.

2. The Asserted Claims

Complainants assert claims 1, 4, 7, and 8 of the '877 patent, which are the patent's four independent claims. Claims 1 and 7 relate to the decoding process and recite similar limitation in different form—claim 1 is directed to a *decoding method* with a series of steps, and claim 7 is directed to a *decoding device* with a circuit configured to perform various functions:

Claim 1	Claim 7
1. A decoding method of a binary stream to reconstruct a current block of a current image from a reference block of a reference image reconstructed at a different size from the size of said current image, said reconstructed reference image being stored in a decoded picture buffer comprising: motion compensating said reference block of said reconstructed reference image by applying a single horizontal filter GFH and a single vertical filter GFv successively on the lines and on the columns of pixels of said reference block, decoding, for the current block, a residue block, and reconstructing the current block from said residue block and from said motion compensated reference block, wherein said single vertical filter GFv applied on a pixel s is such that $GFv(s)=MCIFv(SCFv(s))$, where MCIFv is a vertical motion compensation interpolation filter and SCFv is a vertical resampling filter,	7. A decoding device comprising at least one circuit configured to: access, from a decoded picture buffer, a reference image reconstructed at a size different from the size of a current image; motion compensate a reference block of said reconstructed reference image by applying a single horizontal filter GFH and a single vertical filter GFv successively on the lines and on the columns of pixels of said reference block, decode, for a current block of said current image, a residue block from a binary stream, and reconstruct the current block from said residue block and from said motion compensated reference block, wherein said single vertical filter GFv applied on a pixel s is such that $GFv(s)=MCIFv(SCFv(s))$, where MCIFv is a vertical motion compensation interpolation filter and

MCIFv and SCFv being applied jointly and wherein said single horizontal filter GFH applied on a pixel u is such that $GFH(u)=MCIFH(SCFH(u))$, where MCIFH is a horizontal motion compensation interpolation filter and SCFH is a horizontal resampling filter, MCIFH and SCFH being applied jointly and wherein no resampled version of said reconstructed reference image is stored in the decoded picture buffer.	SCFv is a vertical resampling filter, MCIFv and SCFv being applied jointly and wherein said single horizontal filter GFH applied on a pixel u is such that $GFH(u)=MCIFH(SCFH(u))$, where MCIFH is a horizontal motion compensation interpolation filter and SCFH is a horizontal resampling filter, MCIFH and SCFH being applied jointly and wherein no resampled version of said reconstructed reference image is stored in the decoded picture buffer.
---	---

JXM-1 ('877 patent) at claims 1, 7.

Claims 4 and 8 relate to the encoding process and, as with claims 1 and 7, recite similar limitations in different form. Claim 4 is directed to a *coding method* with a series of steps, and claim 8 is directed to a *coding device* with a circuit configured to perform various functions:

Claim 4	Claim 8
4. A coding method of a current block of a current image from a reference block of a reference image reconstructed at a different size from the size of said current image, said reconstructed reference image being stored in a decoded picture buffer comprising: motion compensating said reference block of said reconstructed reference image by applying a single horizontal filter GFH and a single vertical filter GFv successively on the lines and on the columns of pixels of said reference block, calculating, for the current block, a residue block from said current block and from said motion compensated reference block, and coding the residue block in a binary stream,	8. A coding device comprising at least one circuit configured to: access, from a decoded picture buffer, a reference image reconstructed at a size different from the size of a current image; motion compensate a reference block of said reconstructed reference image by applying a single horizontal filter GFH and a single vertical filter GFv successively on the lines and on the columns of pixels of said reference block, calculate, for a current block of said current image, a residue block from said current block and from said motion compensated reference block, and code the residue block in a binary stream, wherein said single vertical filter GFv applied on a pixel s is such that

wherein said single vertical filter GFv applied on a pixel s is such that $GFv(s)=MCIFv(SCFv(s))$, where MCIFv is a vertical motion compensation interpolation filter and SCFv is a vertical resampling filter, MCIFv and SCFv being applied jointly and wherein said single horizontal filter GFH applied on a pixel u is such that $GFH(u)=MCIFH(SCFH(u))$, where MCIFH is a horizontal motion compensation interpolation filter and SCFH is a horizontal resampling filter, MCIFH and SCFH being applied jointly and wherein no resampled version of said reconstructed reference image is stored in the decoded picture buffer.	$GFv(s)=MCIFv(SCFv(s))$, where MCIFv is a vertical motion compensation interpolation filter and SCFv is a vertical resampling filter, MCIFv and SCFv being applied jointly and wherein said single horizontal filter GFH applied on a pixel u is such that $GFH(u)=MCIFH(SCFH(u))$, where MCIFH is a horizontal motion compensation interpolation filter and SCFH is a horizontal resampling filter, MCIFH and SCFH being applied jointly and wherein no resampled version of said reconstructed reference image is stored in the decoded picture buffer.
--	--

Id. at claims 4, 8.

B. '859 Patent

1. Entropy Encoding Background

The '859 patent relates to a video compression encoding and decoding technique called “entropy coding.” When a video is encoded, it is first transformed into “symbols” that represent the video signal. An entropy encoder then converts the symbols into binary codes that take up less space. JXM-2 ('859 Patent) at 1:23–33. These binary codes make up the compressed bitstream, which is more efficient to store and transmit than the symbols themselves.

To convert symbols into binary codes, the entropy coder needs to have a prediction of how often each symbol is likely to appear (i.e., each symbol’s “probability” of appearing) in the data to be encoded. *Id.* The '859 patent calls this collection of probabilities a “probability model.” *Id.* at 4:39–43. Using the probability model, the entropy encoder can assign smaller binary codes (e.g., ‘0’ or ‘1’) to symbols that are likely to appear more often and larger binary codes (e.g., ‘10110110’

or ‘10110111’) to symbols that are likely to appear less often. *Id.* That way, the symbols likely appear the most often can be converted to binary codes that take up the least amount of space, which provides compression.

As an analogy, consider encoding a phonebook for the Washington, DC area. The last name “Smith” is likely to appear more often than “Stephanopoulos.” RXM-41 (Orchard Decl.) at ¶ 31. Knowing this, an entropy encoder could assign “Smith” a binary code of ‘0’ and “Stephanopoulos” a binary code of ‘10110110.’ *Id.* When encoding the phonebook, the entropy encoder would convert thousands of “Smith” names to the value ‘0’ (which takes up one bit), whereas it would convert perhaps only a handful of “Stephanopoulos” names to the value ‘10110110’ (which takes up eight bits). *Id.* By assigning the highest probability name (“Smith”) a small binary code (one bit), the entropy encoder saves space in the encoded bitstream. *Id.*

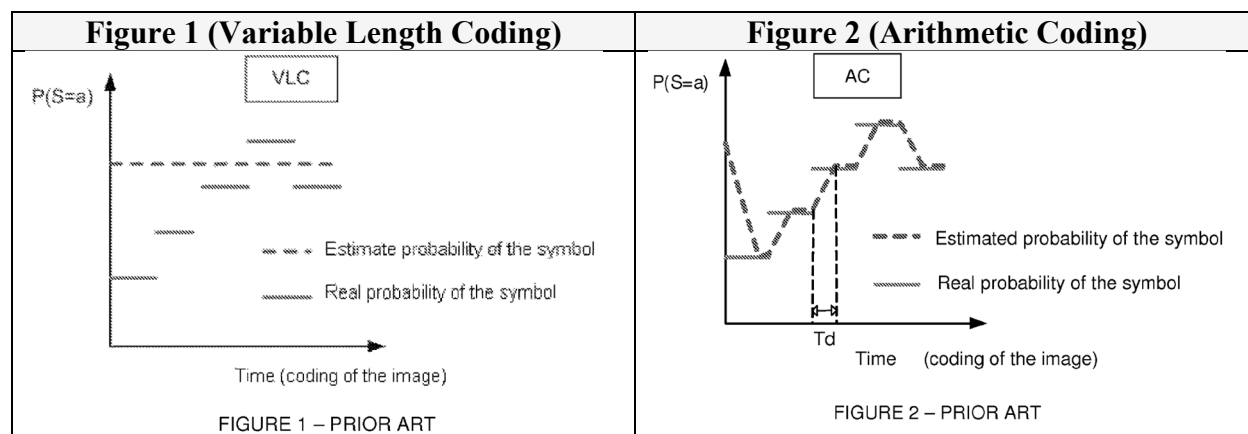
The efficiency of the entropy encoder (i.e., how much compression it can achieve) depends on how well its “probability model” matches the *actual* rate each symbol appears in the data to be encoded: the better the match, the better the efficiency. *See* JXM-2 (’859 Patent) at 2:29–34. The ’859 patent purportedly addresses *one* way to improve the match between the probability model and the data to be encoded, at any given time. *Id.*

2. Admitted Entropy Coding Prior Art

The ’859 patent does not claim to have invented entropy coding. *See, e.g., id.* at 1:23–25 (“To transform symbols representative of a signal, for example a video or audio signal, into binary codes, *it is known in the art* to use a method of entropy coding.” (emphasis added)). The ’859 patent discusses two prior art examples. First, *Variable Length Coding* uses a probability model defined in advance of encoding “by analysing a very large number of signals representative of signal that must be coded subsequently.” *Id.* at 1:41–42. The disadvantage of that approach is that

it uses a “constant” probability for each symbol that can “be distanced from the real probability” of the data being encoded at any given point in time. *Id.* at 1:65–2:1, 2:4–7. Second, *Arithmetic Coding* uses probabilities that are “updated regularly during the image coding of the image slice in order to take into account the characteristics of the source signal that is being coded.” *Id.* at 1:51–54. That is, the probability model is updated as encoding proceeds and the encoder learns more about how often each symbol actually appears in the source signal. Nevertheless, there is still a “delay time” before the estimated probability of each symbol gets “closer to the real probability” of the source signal, since the encoder needs to learn about the real probability of each symbol and adapt to it as encoding proceeds. *Id.* at 2:4–11.

The '859 patent illustrates the alleged shortcomings of these prior art techniques in Figures 1 and 2, shown below. According to the '859 patent, the estimated probability for a symbol (shown with a dashed line) may not line up very well with its real probability (shown with a solid line) over time as encoding proceeds:



Id. at Figs. 1, 2; *see also id.* at 2:4–11 (“In the two aforementioned examples, the estimated probability is either set a priori according to the analysis carried out previous to the coding of the current source signal and can thus be distanced from the real probability (VLC case), or it is regularly updated during the coding of the current signal in order to get closer to the real probability

but with a delay time (AC case) all the greater as the real probability fluctuates during the coding process.”). According to the ’859 patent, this leads to an inefficient encoding/decoding process. *Id.* at 2:29–34.

3. Alleged Invention of the ’859 Patent

The ’859 patent claims to improve prior art entropy coding techniques by introducing “switching points” into the stream during the encoding process. *Id.* at 2:23–33. At these “switching points,” the encoder can select a probability model, from a number of different options, that is likely to be the best match for the source signal. *Id.* From there, the encoder can also update the probability model over time (as with the prior art Arithmetic Coding) as encoding proceeds and the encoder learns additional information about the real probability of the symbols in the source signal. *Id.* According to the ’859 patent, this ability to periodically select a different probability model at “switching points” allows the probability model to more closely track the real probability of the source signal over time. *Id.* The ’859 patent shows this alleged benefit in Figure 9:

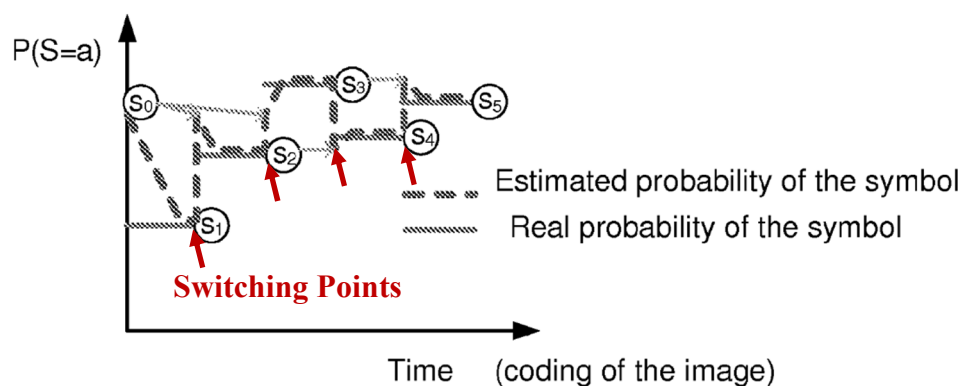
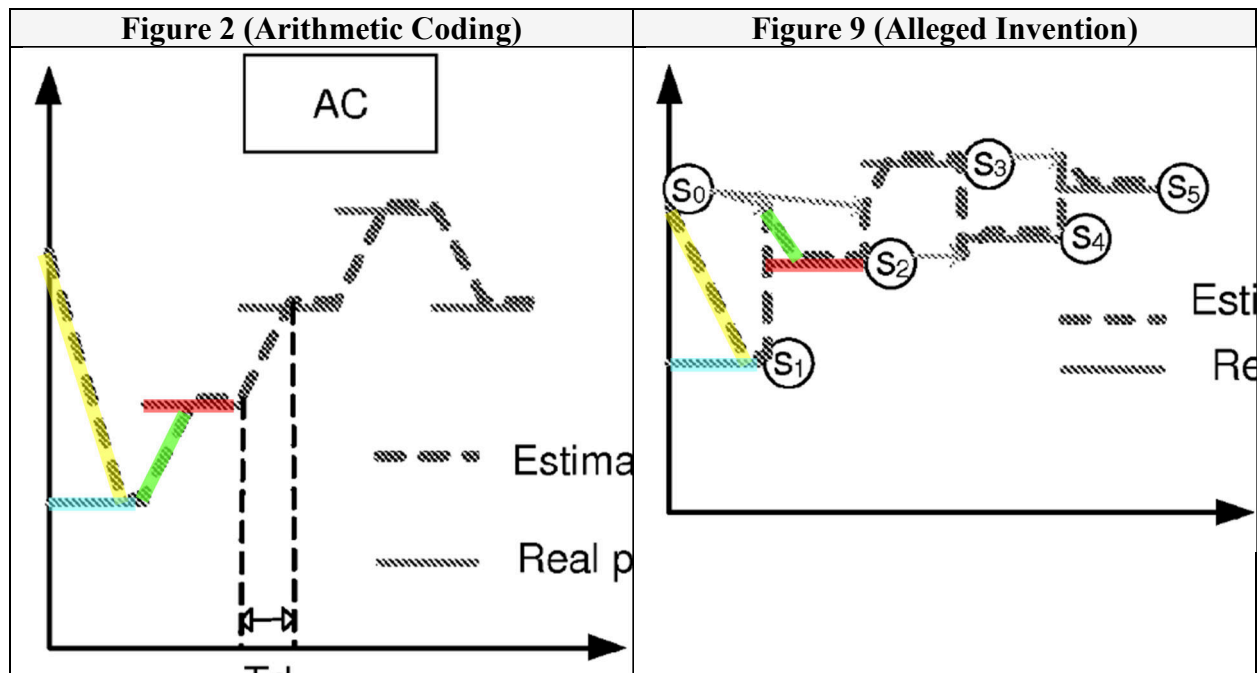


FIGURE 9

Id. at Fig. 9 (annotated); *see also id.* at 9:46–49 (“The full lines represent the change in the real probability $P(S=a)$ and the dotted lines represent the corresponding estimated probability during

the coding of the image according to the method of the invention”). As shown, the estimated probability for any given symbol (shown with a dashed line) is periodically changed at switching points when a different probability model is selected. From there, the probability model is updated over time as encoding proceeds. The ’859 patent claims that this allows for each symbol’s estimated probability to more closely match its real probability over time, when compared to the prior art Arithmetic Coding method. *Id.* at 9:53–54. (“The convergence time is considerably reduced with respect to the AC method (FIG. 2).”)

The claimed improvement resulting from “switching points” can be seen visually, by comparing the prior art Arithmetic Coding method shown in Figure 2 with the alleged invention shown in Figure 9:



Id. at Figs. 2 & 9 (annotated). As shown, both Arithmetic Coding and the alleged invention begin with a probability model and update the model over time as encoding proceeds, so that the estimated probability for any given symbol (shown in yellow) eventually begins to converge on its real probability (shown in blue). The difference is what happens next. According to the ’859

patent, Arithmetic Coding continues to use the same probability model, while updating it over time as encoding proceeds. *Id.* at 1:47–54, 1:67–2:3. As a result, when the real probability for any given symbol changes in the encoding process, it again takes some time for the estimated probability (shown in green) to converge on the real probability (shown in red) of the symbol in the source signal. *Id.* at 2:7–11.

To improve this convergence time, the '859 patent takes advantage of “switching points” to select a different probability model at various points in the encoding process—one that is likely to be a better match for the source signal at that particular point than the probability model currently in use. *Id.* at 2:23–33. In Figure 9 above, the encoder changes the probability model at the first “switching point,” which results in changing the estimated probability for the symbol. Then, as with Arithmetic Coding, the alleged invention of the '859 patent continues to update the probability model over time, such that the estimated the estimated probability (shown in green) converges on the real probability (shown in red) of the source signal. Compared to the Arithmetic Coding example shown in Figure 2, the alleged invention supposedly allows for faster convergence because the “switching points” allows the encoder to select a probability model that is a closer estimate of the source signal than the probability model being used at that point in the encoding process. *Id.* at 2:29–33, 9:51–54; *see also* RXM-41 (Orchard Decl.) at ¶¶ 40–46.

The '859 patent further describes the “switching points” in connection with the encoding process outlined in Figure 3 (shown at right). When encoding a symbol, at step 100 the encoder first determines whether it is “situated on a switching point,” which the patent describes as “a point in the stream at the level of which the probability model used to code the current symbol can be modified.” *Id.* at Fig. 3 (step 100), 4:56–60. If the encoder is *not* at a “switching point,” it encodes the current symbol (S_c) with the current probability model (step 106) and then updates the current probability model to account for the symbol it just encoded (step 108). *Id.* at 5:8–9, 5:42–46. According to the '859 patent, these are the same steps already known in Arithmetic Coding. *See id.* at 5:44–46 (“During a step 108, the current probability model P_c is updated according to the method described previously in reference to the AC coding.”). In short, if *not* at a switching point, the encoder simply uses the approach already known in the Arithmetic Coding—it encodes the symbol and updates the probability model.

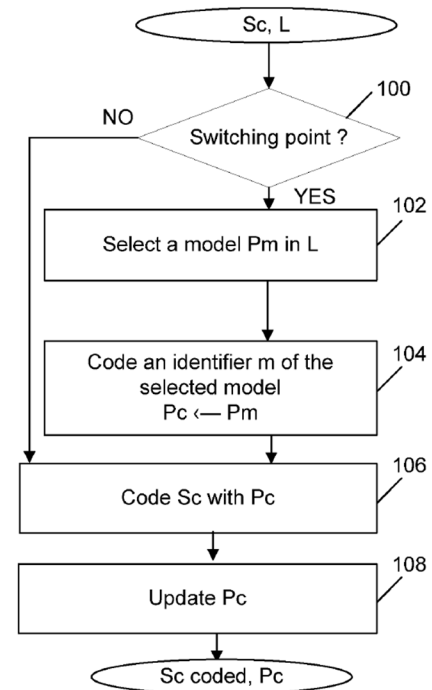


FIGURE 3

If the encoder determines the current symbol *is* located at a “switching point,” it performs two additional steps according to the invention. First, the encoder selects a probability model (P_m) to use going forward, from among the set of probability models (L) that it has available (step 104). *Id.* at 5:10–14.² Then, the encoder codes an “identifier” of the selected probability model (P_m)

² The '859 patent purports to describe how this selection is made, i.e., how the encoder determines which probability model would be optimal according to “cost criterion,” *id.* at 5:15–35, although the details are beyond the scope of this brief.

into the stream and makes the selected probability model (P_m) the current probability model (P_c) to use going forward (step 106). *Id.* at 5:36–43. The “identifier” is coded into stream because ultimately, a decoder will need to decode the encoded bitstream to play back the video, and to do that, the decoder needs to know which probability model the stream was encoded with. *Id.* at 5:37–39. The encoder then encodes the current symbol and updates the current probability model using the same Arithmetic Coding steps (106 and 108) previously discussed. *Id.* at 5:42–46. Finally, the encoder repeats the steps shown in Figure 3 for each symbol, as it continues to encode the video. *Id.*; see also RXM-41 (Orchard Decl.) at ¶¶ 47–49.

The '859 patent describes a similar series of steps for *decoding* an encoded video in connection with Figure 11. When decoding a symbol, the decoder first determines whether it is located at a “switching point” (step 200). *Id.* at 10:64–11:9. If the decoder is not at a “switching point,” it operates the same way as Arithmetic Coding. *Id.* at 11:9–11. That is, the decoder decodes the current symbol (S_c) with the current probability model (step 206) and then updates the current probability model to account for the symbol it just decoded (step 208). *Id.* at 11:55–59. Here again, the '859 patent concedes that the steps the decoder performs when *not* at a switching point were already known and used in Arithmetic Coding. See *id.* at 11:57–59 (“During a step 208, the current probability model P_c is updated according to the method described previously in reference to the AC coding.”).

If the decoder determines the current symbol *is* located at a “switching point,” it performs three additional steps according to the alleged invention. First, the decoder decodes an “identifier”

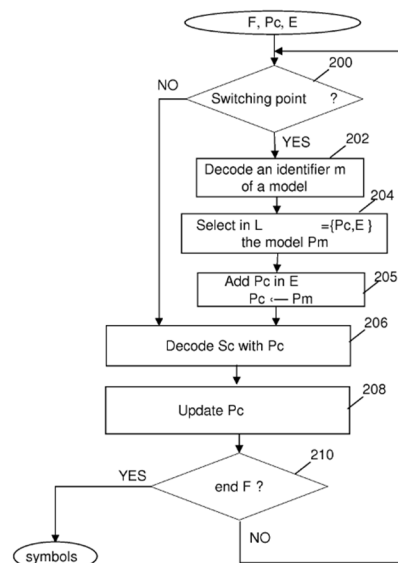


FIGURE 11

from the stream that informs the decoder which probability model to use (step 202). *Id.* at 11:11-17. Second, the decoder then uses the “identifier” to select a probability model (Pm), from among a set of probability models (L) that it has available (step 204). *Id.* at 11:11-17. Third, the decoder “adds” the current probability model (Pc) into a subset of available probability models and makes the selected probability model (Pm) the current probability model (Pc) to use going forward (step 205). The decoder then decodes the current symbol and updates the current probability model using the same Arithmetic Coding steps (206 and 208) previously discussed. *Id.* at 11:55-61. Finally, the decoder repeats the steps shown in Figure 11 for each symbol, as it continues to decode the video. *Id.* at 11:61-64; *see also* RXM-41 (Orchard Decl.) at ¶¶ 50-51.

4. The Asserted Claims of the '859 Patent

Complainants assert claims 10 and 15 of the '859 patent, which relate to the decoding process. The claims recite similar limitations in different forms—claim 10 is directed to a *method* with a series of steps, and claim 15 is directed to a *decoder* with a series of means-plus-function limitations:

Claim 10	Claim 15
10. Arithmetic decoding method for symbols coded in the form of a stream, comprising the following steps applied at switching points distributed in said stream to decode a current symbol: decoding from the stream, for said current symbol, a probability model identifier, selecting using said probability model identifier a probability model in a set comprising a current probability model defined from symbols coded previous to the current symbol and a subset of probability models wherein subset comprises at least one probability model defined by default,	15. Arithmetic decoder for decoding symbols coded in the form of a stream, comprising: means for decoding from the stream, at switching points distributed in said stream to decode a current symbol, for said current symbol, a probability model identifier, means for selecting using said probability model identifier a probability model in a set comprising a current probability model defined from symbols coded previous to the current symbol and a subset of probability models wherein subset comprises at least one probability model defined by default,

adding in said subset said current probability model, the selected probability model becoming the current probability model, decoding said current symbol with said current probability model, and updating said current probability model according to the decoding of said current symbol.	means for adding in said subset said current probability model, the selected probability model becoming the current probability model, means for decoding said current symbol with said current probability model, and means for updating said current probability model according to the decoding of said current symbol.
--	--

JXM-2 ('859 Patent) at claims 10, 15.

C. '556 Patent

1. Background

The '556 patent relates to filters applied during the video encoding and decoding process, specifically focused on reducing visual defects known as “artifacts” or “quantization artifacts.” JXM-3 ('556 Patent) at 4:58–67. Artifacts result from a process called “quantization,” a key step in the video compression process. *Id.* at 1:12–14. There are several types of quantization artifacts, including block artifacts, ringing artifacts, and “independent and identically distributed (i.i.d.) noise.” *Id.* at 1:14–17, 2:25–30. The '556 patent concedes that conventional filters for addressing these types of artifacts were well known, such as deblocking filters for block artifacts, deringing filters for ringing artifacts, and sparse de-noising filters for independent and identically distributed (i.i.d.) noise. *Id.* at 1:23–32, 2:04–11; 10:13–16; 16:43–47.

The purpose of a “deblocking” filter is to smooth the edges around each decoded block, such that the individual blocks no longer appear discontinuous. Such “deblocking” filters were known in the prior art. As an example, the '556 patent recognizes that the prior art MPEG-4 AVC standard used a well-known adaptive in-loop deblocking filter that is effective at reducing blocky artifacts, in part because of its ability to adapt the strength of its filtering by the values of several syntax elements. JXM-3 ('556 Patent) at 1:33–38. Another benefit is its ability to distinguish

between a true edge (i.e., an edge that originally existed in the image) from a “block edge” (i.e., a visual edge or block that is a result of quantization and appears to the viewer as distortion of the image), and only smooth out the latter. *Id.* at 1:38–46. While known deblocking filtering methods are effective at reducing blocky artifacts, the ’556 patent claims they are not effective at correcting other artifacts. *Id.* at 1:58–61.

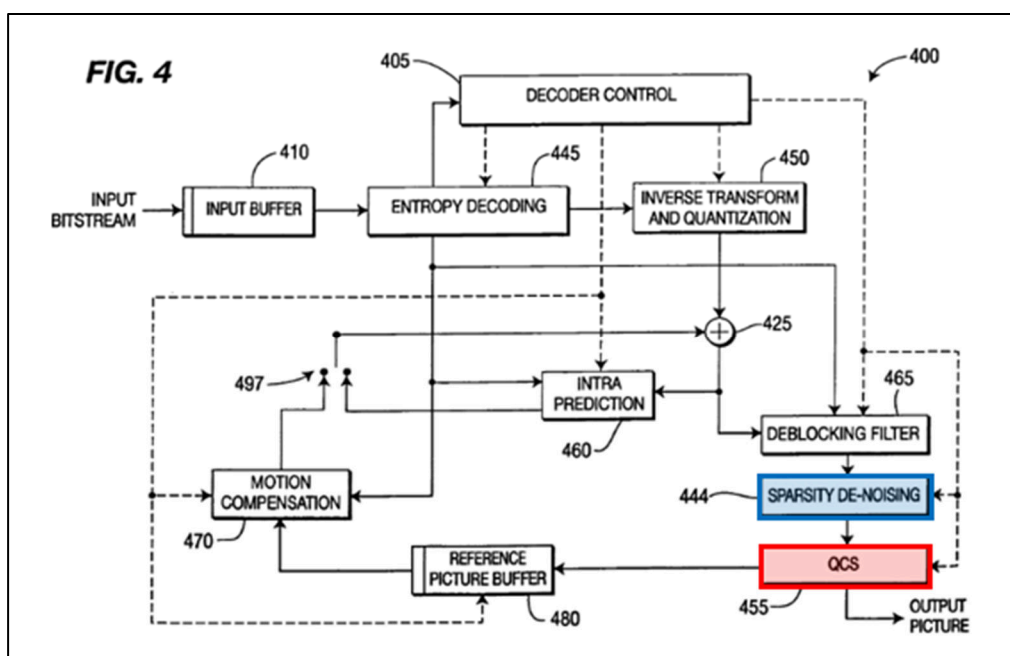
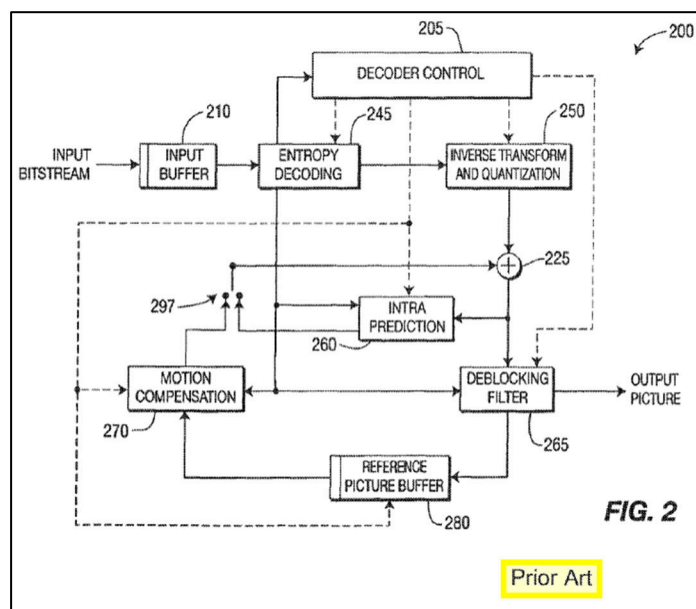
The ’556 patent also concedes that it was known to use a “sparse de-noising filter” to address a different type of noise called “independent and identically distributed (i.i.d.) noise.” *Id.* at 2:22–30. The “sparse de-noising filter” is a filter “which exploits a sparse image model using an over complete set of linear transforms and hard-thresholding,” as described in greater detail below. JXM-3 (’556 Patent) at 2:08–10. According to the ’556 patent, “adaptive sparse de-noising filtering is directed to recovering the original signal from an independent and identically distributed (i.i.d.) noisy environment.” *Id.* at 14:19–21. Sparse de-noising filters have various benefits, such as reducing quantization noise in “locally uniform regions (smooth, high frequency, texture, and so forth) separated by singularities.” *Id.* at 2:22–25. But they also have known problems, such as a tendency to “get confused by true edges or false block edges.” *Id.* at 2:30–31.

The ’556 patent proposes to apply both filters in succession—deblocking and then sparse de-noising—because “the quantization noise in video compression, and particularly with respect to the MPEG-4 AVC Standard, is usually *not* independent and identically distributed (i.i.d.)” before a deblocking filter is applied. *Id.* at 14:21–24 (emphasis added). The patent claims that “[t]he signal *after* deblocking filter is a better match to the independent and identically distributed (i.i.d.) noise presumption,” making it ripe for filtering by the sparse de-noising filter. *Id.* at 14:28–30 (emphasis added). The ’556 patent claims this particular sequence of applying a deblocking filter followed by a sparse de-noising filter was not known in the prior art. *Id.* at 14:24–28 (“The

present principles differ from the prior art in the use of a *series of refinement steps (succession of filters)* after, for example, the MPEG-4 AVC Standard deblocking filter, each oriented to correct at least one part of the coding artifacts.” (emphasis added)).

At the same time, concerned that the subsequent application of a sparse de-noising filter can over-smooth a textured region of the image (e.g., where the region is not uniform but instead displays a texture, such as a carpet), the ’556 patent proposes applying a third filter known as the QCS filter. *Id.* at 14:30–34 (“The possible “over filtering problem” resulting from the use and/or combination of deblocking and de-noising filters is solved by the subsequent quantization constraint set (QCS) filter.”). The QCS is an additional filter applied to control for “over-smoothing” caused by the sparse de-noising filter. *Id.* at 11:11–15. The result is a multi-pass filter, which the ’556 patent claims to improve quantization noise reduction in video frames. *Id.* at 14:34–36.

A comparison of Figures 2 and 4 is helpful to understand the ’556 patent’s purported change to the prior art. Figure 2 shows prior art: a block diagram for a video decoder according to the known MPEG-4 AVC standard. *Id.* at 5:31–34. Figure 4 shows an embodiment of the claims of the ’556 patent. *Id.* at 5:39–43. The only difference between the two figures is the ’556 patent’s introduction of a “sparse de-noising filter” (**blue**) and a “QCS” or “quantization constraint set” (**red**) after the known deblocking filter.



JXM-3 ('556 Patent) at Figs. 2, 4 (annotations added).

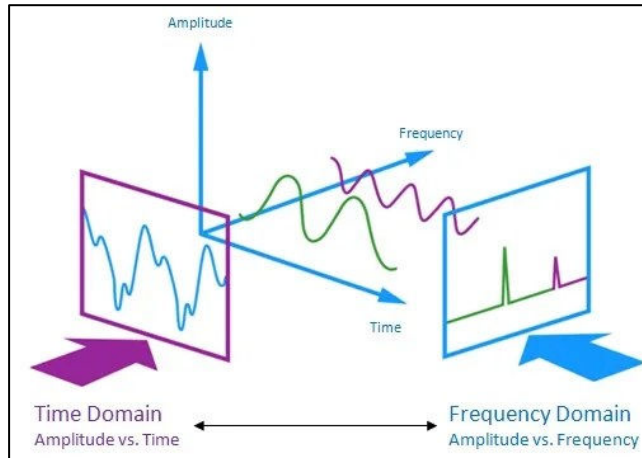
Relevant here, in *addition* to deblocking and sparse de-noising filters, the specification of '556 patent also lists “deringing” filters as a type of filter that targets quantization artifacts. JXM-3 ('556 Patent) at 10:13–16. Deringing filters are designed to filter ringing artifacts, which are

“[r]ipple-like artifacts around sharp edges in a decoded image.” RXM-26 (Richardson, The H.264 Advanced Video Compression Standard) at xix. The ’556 patent does not claim deringing filters.

2. Sparse De-noising Filters

As noted above, the ’556 patent describes a “sparse de-noising filter” as a filter “which exploits a sparse image model using an over complete set of linear transforms and hard-thresholding.” *Id.* at 2:08–10. The “approach basically includes three steps: *transform*; transform coefficients *threshold*, and *inverse transform*.” ’556 Patent, at 2:14–16. That is, the image begins in the “pixel domain.” RXM-41 (Orchard Decl.) at ¶ 63. A transform is applied to the image, and the result is a representation of the original image in the “transform domain.” *Id.* at ¶ 63. Then, all values in the representation below a certain threshold value (which represent noise) are filtered out, while values above the threshold value (which represent the true signal) are kept. *Id.* at ¶ 65. Finally, the image is reconstructed by applying an inverse transform. *Id.* at ¶ 66.

As an analogy, consider a piano concert where a pianist plays a complex chord, but there is faint background noise, like a whispering crowd. *Id.* at ¶ 69. In the “time domain,” the audio signal is represented by a complex wave. *Id.* Using a technique like the Fourier transform, the complex sound wave of the chord (and the whispers) is transformed into the frequency domain. *Id.* In the frequency domain, the chord becomes a set of distinct peaks, each representing a note, while the background noise appears as minor fluctuations. *Id.* This transformation is illustrated below, where the complex wave in the time domain is shown at left, and the result in the frequency domain is shown at right. *Id.*

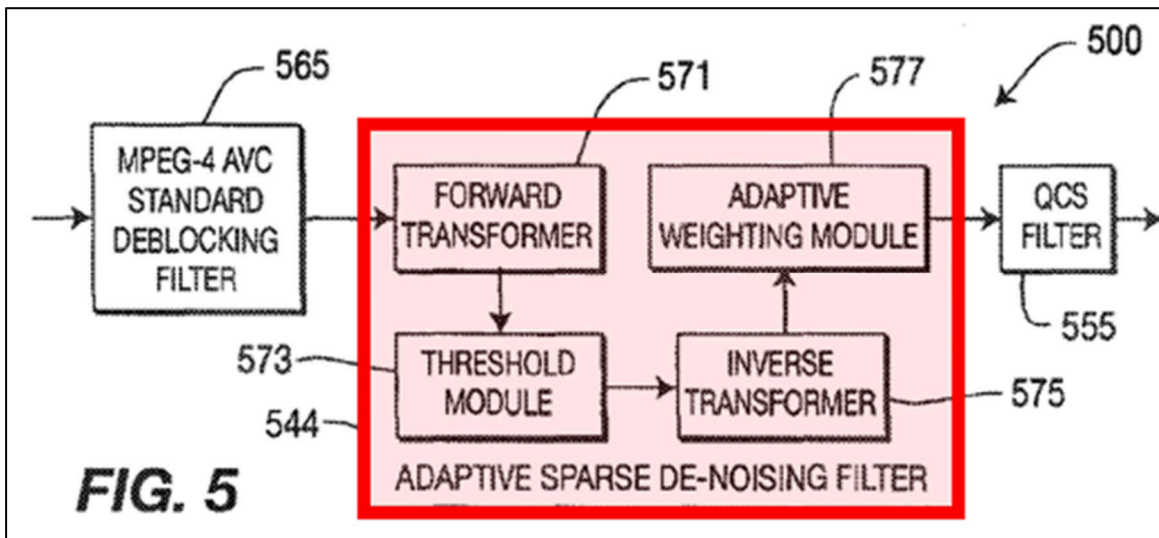


<https://play.fallows.ca/wp/radio/ham-radio/signal-analysis-morse-decoder/>

In the frequency domain, the significant notes of the chord (shown in green and purple) can be differentiated from the noise. *Id.* at ¶ 71. By setting a threshold, the elements below it (representing the background noise) can be filtered out. *Id.* After applying an inverse transform, the result is a clearer sound wave of the original chord, with the noise filtered out. *Id.*

In video processing, a decoded image or frame may contain “independent and identically distributed (i.i.d.) noise,” analogous here to the faint whispers of the crowd. *Id.* at ¶ 72. To correct for i.i.d. noise, a sparse de-noising filter applies three steps. *Id.* First, a transform is applied to the image into the “transform domain.” *Id.* at ¶¶ 60–64.³ Then, a threshold module is applied to eliminate the noise but retain the prominent natural features of the image. *Id.* at ¶ 65. Finally, an inverse transform is applied to reconstruct the original image, free of any i.i.d. noise, in the “pixel domain.” *Id.* at ¶ 66. This process is illustrated in Figure 5:

³ In the context of video processing, typical transforms include the Discrete Cosine Transform (DCT), the Wavelet Transform, and the Fast Fourier Transform (FFT). *Id.* at ¶ 62.

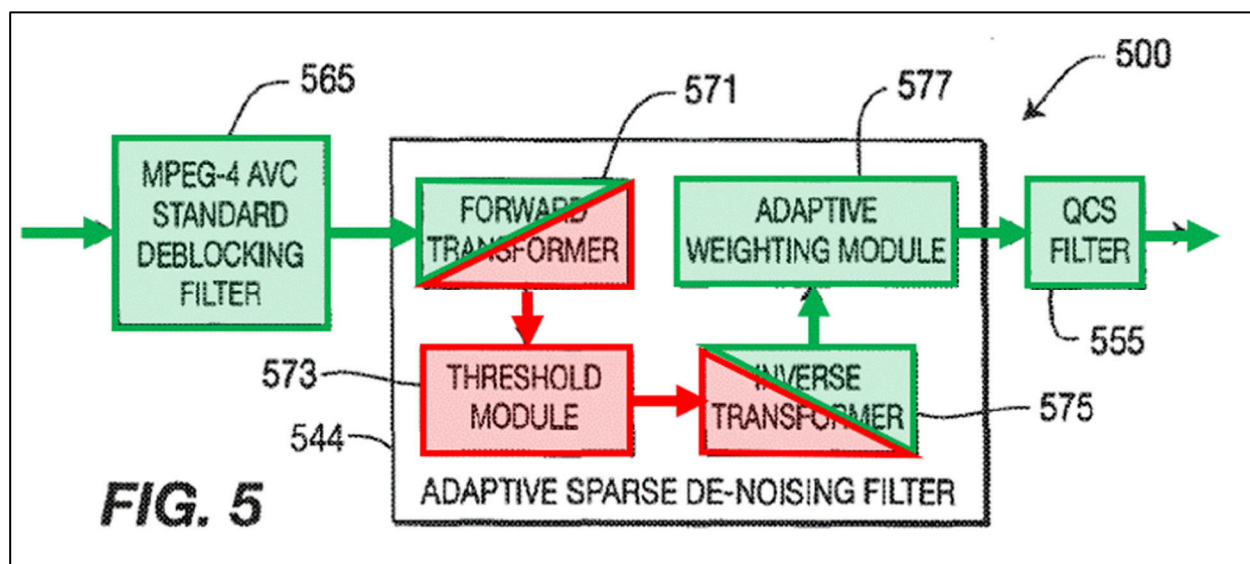


JXM-3 ('556 Patent) at Fig. 5 (annotations added).

The term “sparse” in the context of video processing generally refers to a representation of an image with “few significant non-zero coefficients.” See RXM-34 (Ravishankar, Learning Sparsifying Transforms) at 1 (“It is well known that natural signals and images have an essentially *sparse representation (few significant non-zero coefficients)*.” (emphasis added)); RXM-35 (U.S. Pat. App. 2011/0222597) at [0022] (“A sparse matrix is a matrix *populated primarily with zeroes*.” (emphasis added)). Images are not sparse in the *pixel* domain, because each pixel has a value. RXM-41 (Orchard Decl.) at ¶ 68. The *representation* of an image in the *transform* domain (after a transform has been applied), however, is typically sparse. *Id.* That is, there will be a few non-zero values representing the natural features of the image, but all other values in the matrix are zero (or, where noise is present, near-zero). *Id.* Using the threshold module, all near-zero values (the i.i.d. noise) can be filtered out, while retaining the larger non-zero values (the natural features of the original image). *Id.* at ¶ 65.

Importantly, the image itself is not “sparse”—the underlying data only becomes “sparse” *in the transform domain*, that is, after a transform has been applied. *Id.* at ¶ 68. Likewise, after an inverse transform is applied, and the data is returned to the pixel domain, the data is no longer

“sparse.” *Id.* To illustrate, the underlying data comprising the image is “sparse” during the process steps in **red**, the image is not sparse during the process steps in **green**. *Id.*



Id. (annotating JXM-3 (’556 Patent) at Fig. 5).

The transforms applied by the sparse de-noising filter are described as “over complete” because the filter does not take just a *single* transformed, de-noised sample. RXM-41 (Orchard Decl.) at ¶ 73. Rather, the patent describes that “*several* de-noised estimates provided by de-noising with an over complete set of transforms *are combined* using weighted averaging at every pixel.” JXM-3 (’556 Patent) at 2:16–21 (emphases added). This combination of the several de-noised estimates (made using an over complete set of transforms) occurs within the “adaptive weighting module” 577 of Figure 5, shown above. RXM-41 (Orchard Decl.) at ¶ 73.

3. Alleged Invention of the ’556 Patent

The ’556 patent does not claim to have invented filtering techniques like deblocking filters or sparse de-noising filters. The patent concedes that these types of filters were already known in the field of video coding. JXM-3 (’556 Patent) at 1:33–38, 2:04–10. The claimed novelty of the ’556 patent lies in the application and combination of these existing filters. The patent specifically

details a sequential use of these filters, where a deblocking filter is employed first to reduce blocky artifacts, followed by a sparse de-noising filter to further reduce quantization noise:

The present principles differ from the prior art in the use of a series of refinement steps (succession of filters) after, for example, the MPEG-4 AVC Standard deblocking filter, each oriented to correct at least one part of the coding artifacts. The signal after deblocking filter is a better match to the independent and identically distributed (i.i.d.) noise presumption. The possible “over filtering problem” resulting from the use and/or combination of deblock and de-noising filters is solved by the subsequent quantization constraint set (QCS) filter. The succession of filters has shown its superior performance in removing video compression artifacts, both in practice and theory.

JXM-3 ('556 Patent) at 14:24–36 (emphases added).

The individual filters themselves, however, are well known in the art, and the '556 patent does not claim to have invented them.

4. The Asserted Claims of the '556 Patent

Complainants assert claims 1, 3, 4, 5, 7, and 8 of the '556 patent, which relate to the decoding process. The independent claims are claims 1 and 5. They recite similar limitations in different forms—claim 1 is directed to a *method* with a series of steps, and claim 5 is directed to an *apparatus* configured to perform equivalent steps:

Claim 1	Claim 5
1. A method comprising: decoding an image region, wherein the decoding includes performing in-loop filtering to reduce at least a first and a second type of coding artifact using at least two filters in succession, the at least two filters including a deblocking filter for performing a first pass to reduce blocking artifacts	5. An apparatus comprising at least a memory and one or more processors, the one or more processors being configured to: decode an image region, the one or more processors including at least two filters for successively performing in-loop filtering to reduce at least a first and a second type of coding artifact, the at least two filters including a deblocking filter for performing a first pass to reduce blocking artifacts

and an adaptive sparse de-noising filter for performing a second pass to reduce noise, wherein the adaptive sparse de-noising filter is selectively enabled or disabled at a given level, the given level being at least a macroblock level, a slice level, a picture level or a sequence level.	and an adaptive sparse de-noising filter for performing a second pass to reduce noise, wherein the adaptive sparse de-noising filter is selectively enabled or disabled at a given level, the given level being at least a macroblock level, a slice level, a picture level or a sequence level.
--	--

JXM-3 ('556 patent) at claims 1, 5.

IV. LEVEL OF ORDINARY SKILL IN THE ART

The parties have agreed that a POSITA for the '877, '859, and '556 patents would have had “at least a bachelor’s degree in electrical Engineering, computer science, computer engineering, or a related field, and approximately three years of experience in design, development, and/or testing of relevant multimedia processing, such as video, and/or image processing or coding. Advanced graduate studies with coursework and/or research focusing on topics of video encoding and decoding could substitute for working experience in the field.”

V. '877 PATENT

A. “circuit configured to” (claims 7, 8)

Complainants’ Construction	Respondents’ Construction	Staff’s Construction
Plain and ordinary meaning (i.e., requires pre-existing hardware and software programmed to perform the cited functionality)	Plain and ordinary meaning (i.e., requires pre-existing programming of hardware and software to perform the cited functionality)	Plain and ordinary meaning (i.e., requires pre-existing programming of hardware and software to perform the cited functionality)

The parties agree that “circuit configured to” should be given its plain and ordinary meaning, but dispute what the plain and ordinary meaning of the term is. Respondents and Staff agree the plain and ordinary meaning is “requires pre-existing programming of hardware and

software to perform the cited functionality.” Complainants argue the plain and ordinary meaning is “requires pre-existing hardware and software programmed to perform the cited functionality.” Respondents respectfully request the ALJ adopt Staff’s and Respondents’ proposed construction.

When the parties filed their Joint Claim Construction Chart (“JCCC”) on December 13, 2023, the dispute surrounding “circuit configured to” centered on whether that language is limiting in the context of claims 7 and 8 of the ’877 patent. EDIS Doc. Id. 810282 (Dec. 13, 2023 JCCC) at 2. The parties agreed that, if “circuit configured to” is limiting, it should be given its plain and ordinary meaning, with minor differences in the proposed language expressing the plain and ordinary meaning of the term.⁴ *Id.* Through the meet and confer process, the parties came to agreement as to the plain and ordinary meaning of “circuit configured to”—i.e., “requires pre-existing programming of hardware and software to perform the recited functionality.”

On the January 18, 2024 Discovery Committee Meeting (“DCM”), Complainants inexplicably changed their position, agreeing that, for the purposes of this Investigation, “circuit configured to” is limiting in claims 7 and 8 of the ’877 patent, and proposing a different phrasing of the plain and ordinary meaning as reflected in the chart above. Complainants have provided no explanation for their alternate language or how the changes they made are allegedly different than what they had previously agreed to.

The term “‘configured to’ is a patent term of art used in many fields of art” that has a narrower meaning than “capable of.” *Certain Computer Network Security Equip. and Systems*,

⁴ Staff and Complainants’ JCCC proposals were aligned and matched that of Respondents and Staff as of this brief—“i.e., requires pre-existing programming of hardware and software to perform the recited functionality.” EDIS Doc. Id. 810282 (Dec. 13, 2023 JCCC) at 2. Respondents’ JCCC proposal was nearly identical but added the words “without modification” to the end—i.e., “requires pre-existing programming of hardware and software to perform the recited functionality *without modification*.” *Id.*

Inv. No. 337-TA-1314, Order No. 37 at 11-12 (USITC Feb. 22, 2023) (McNamara, ALJ) (citing *Radware Ltd. v. A10 Networks, Inc.*, No. C-13-02024-RMW, 2014 WL 1572644, at *12 (N.D. Cal. Apr. 18, 2014)); *Aspex Eyewear, Inc. v. Marchon Eyewear, Inc.*, 672 F.3d 1335, 1348-49 (Fed. Cir. 2012) (noting “configured to” is narrower than “capable of” or “suitable for” while construing “adapted to” as “designed or configured to accomplish the specified objective, not simply that they can be made to serve that purpose”). Respondents’ and Staff’s proposed plain and ordinary meaning is consistent with these principles and with the ALJ’s construction of similar language in *Certain Computer Network Security Equipment*. Inv. No. 337-TA-1314, Order No. 37 at 9-10 (USITC Feb. 22, 2023) (McNamara, ALJ) (construing “configured to” under its “[p]lain and ordinary meaning, i.e., requires pre-existing programming of hardware and software to perform the recited functionality without modification”).⁵ Respondents see no reason to deviate with previously agreed-upon language that is consistent with the ALJ’s construction of similar terms, and Complainants have not even explained what their change to that construction even means.

For the foregoing reasons, Respondents respectfully request the ALJ adopt Respondents’ and Staff’s construction of “circuit configured to”—i.e., “plain and ordinary meaning (i.e., requires pre-existing programming of hardware and software to perform the cited functionality).”

⁵ In response to a question from Staff, Complainants indicated they “are not interpreting ‘configured to’ to include ‘capable of being programmed to perform the recited functionality.’” However, Complainants did not provide any additional detail as to the difference between Staff’s and Respondents’ construction and Complainants’ construction. As such, Respondents are unclear if Complainants’ construction is in keeping with these principles.

VI. '859 PATENT

A. “arithmetic decoding method for symbols coded in the form of a stream, comprising the following steps applied at switching points distributed in said stream to decode a current symbol” (claim 10)

Complainants’ Construction	Respondents’ Construction	Staff’s Construction
The terms “stream” and “current symbol” are limiting at least to the extent that they provide antecedent basis for those terms as they appear elsewhere in the claim. <u>stream</u>: stream of data. <u>current symbol</u>: a symbol that is currently being decoded. “Switching points” is not limiting, but to the extent it is found limiting, a <u>switching point</u> is a point in the stream at the level of which the probability model used to code the current symbol can be modified.	The portion of this claim term following “comprising” is not part of the preamble and therefore is limiting. Regardless, the entire term is limiting. The preamble is limiting. To the extent construction is necessary, <u>stream</u>: “stream of data” <u>current symbol</u>: “a symbol that is currently being decoded” <u>switching point</u>: “a point in the stream at the level of which the probability model used to code the current symbol can be modified—i.e., a probability model can be selected”	The portion of this claim term following “comprising” is not part of the preamble and therefore is limiting. Regardless, the terms “stream” and “current symbol” are limiting at least to the extent that they provide antecedent basis for those terms as they appear elsewhere in the claim. stream: stream of data. current symbol: a symbol that is currently being decoded. The term “switching points” is also limiting, regardless of whether it is considered to be part of the preamble. switching points: a point in the stream at the level of which the probability used to code the current symbol can be modified.

The parties have three disputes about this term: (1) which part is the preamble of claim 10; (2) whether the preamble is limiting; and (3) how “switching points” should be construed.⁶ Each dispute is addressed below.

⁶ The parties have agreed to the construction of “stream” and “current symbol.”

1. The Preamble of Claim 10 is the Language Before “Comprising”

First, the parties dispute which part of the phrase “*arithmetic decoding method for symbols coded in the form of a stream*, comprising the following steps applied at switching points distributed in said stream to decode a current symbol” is the preamble of claim 10. Respondents and Staff agree the preamble is the portion of the term italicized above, which precedes “comprising.” Complainants maintain the preamble is the entire phrase.

Transition terms are used to divide the preamble, which may or may not be limiting, from the claim body, which defines the scope of the claims. *CIAS, Inc. v. Alliance Gaming Corp.*, 504 F.3d 1356, 1360 (Fed. Cir. 2007) (citing 3 Chisum on Patents § 8.06[1][b], at 8–180–82 (2007) for the proposition that “claims usually are structured with a preamble, a ‘transition phrase,’ and the elements or steps that are necessary to the right to exclude”); *Bicon*, 441 F.3d at 949 (“[T]he preamble consist[s] of everything in the claim preceding the word ‘comprising,’”). As the Federal Circuit has explained, “comprising” is a commonly-used transition term. *E.g., id.; In re Fought*, 941 F.3d 1175, 1178 (Fed. Cir. 2019) (“[T]he typical claim language (comprising) . . . denotes the transition between the preamble and the body”); *Gillette*, 405 F.3d at 1371 (“The word ‘comprising’ transition[s] from the preamble to the body”).

Here, the preamble is the phrase “arithmetic decoding method for symbols coded in the form of a stream,” which precedes the transition term “comprising” in claim 10. The phrase “the following steps applied at switching points distributed in said stream to decode a current symbol” that follows “comprising” is located in the body of the claim and is thus limiting (regardless of whether the preamble is also limiting).

Indeed, Complainants do not even contest the related issue in claims 7 and 8 the ’877 patent. There, they agree the phrase “circuit configured to,” which appears after “comprising,”

limits the scope of claims 7 and 8. Likewise, the portion of claim 10 of the '859 patent after “comprising” also limits the scope of that claim because it is not part of the preamble.

2. The Preamble of Claim 10 is Limiting

Regardless of which part the ALJ considers to be the preamble, the entire phrase “arithmetic decoding method for symbols coded in the form of a stream, comprising the following steps applied at switching points distributed in said stream to decode a current symbol” is limiting.

Under Respondents and Staff’s position, the preamble is “arithmetic decoding method for symbols coded in the form of a stream.” As shown below, that phrase provides antecedent basis for “**said stream**” and “**the stream**,” which appear in the claim body:

10. *Arithmetic decoding method for symbols coded in the form of **a stream**, comprising the following steps applied at switching points distributed in **said stream** to decode a current symbol:*

decoding from **the stream**, for said current symbol, a probability model identifier, . . .

JXM-2 ('859 Patent) at claim 10.

One way a preamble becomes limiting is by providing antecedent basis for terms in the claim body. *Catalina*, 289 F.3d at 808. The limiting effect applies to the preamble and is not confined to just the words that provide antecedent basis themselves (since those words are already limitations in the body regardless). The Federal Circuit has recognized a narrow exception, where providing antecedent basis does not necessarily render the entire preamble limiting, but that exception applies only if the preamble can “be neatly packaged into two separate portions,” such that the portion providing antecedent basis (which is limiting) is separate and distinct from the other portion that does not. *Bio-Rad Labs., Inc. v. 10X Genomics Inc.*, 967 F.3d 1353, 1371 (Fed. Cir. 2020). That exception does not apply here. In *Bio-Rad*, for example, the preamble was “[a] method for conducting a *reaction* in plugs in a *microfluidic system*, comprising the steps of.” *Id.*

(emphasis added). The Federal Circuit held the district court erred by finding that the preamble was limiting “only to the extent that it provides an antecedent basis for the terms ‘microfluidic system’ and ‘reaction.’” *Id.* (citation omitted). According to the court, the preamble could not be “splice[d]” into limiting and nonlimiting portions, as the district court had done. *Id.* Instead, the entire preamble was limiting because “[t]he language relied upon for antecedent basis in the preamble at issue is intertwined with the rest of the preamble.” *Id.* The same is true here in claim 10. The word “stream” is intertwined with and cannot “be read separately from,” *id.*, the rest of the phrase “symbols coded in the form of a stream,” because the phrase “symbols coded in the form of” modifies “a stream.” The preamble should be construed as limiting.

Under Complainants’ position that the entire claim term at issue is the preamble, it is limiting for at least two reasons. First, as shown below, the phrase provides antecedent basis for the “**the stream**,” and “**said current symbol**,” and “**the current symbol**,” which appear in the claim body:

10. *Arithmetic decoding method for symbols coded in the form of **a stream**, comprising the following steps applied at switching points distributed in **said stream** to decode **a current symbol**:*

decoding from **the stream**, for **said current symbol**, a probability model identifier,

selecting using said probability model identifier a probability model in a set comprising a current probability model defined from symbols coded previous to **the current symbol** and a subset of probability models wherein subset comprises at least one probability model defined by default,

...

decoding **said current symbol** with said current probability model, and

updating said current probability model according to the decoding of **said current symbol**.

JXM-2 ('859 Patent) at claim 10. For the same reasons discussed above, this antecedent basis renders the entire phrase limiting, because the words “stream” and “current symbol” are intertwined with and cannot be read separately from the remainder of the phrase. The word “stream” specifies where both the coded symbols and “switching points” are located, and the word “current symbol” is part of a phrase making clear that the steps that follow are performed “to decode a current symbol.”

Second, the phrase is also “‘necessary to give life, meaning, and vitality’ to the claim” by reciting that the steps set forth in the body are “applied at switching points” distributed in the stream to decode a current symbol. *Catalina Mktg.*, 289 F.3d at 808 (citation omitted). The determination of whether a preamble is limiting requires a “review of the entire . . . patent to gain an understanding of what the inventors actually invented and intended to encompass by the claim.” *Id.* (quoting *Corning Glass*, 868 F.2d at 1257) (internal brackets omitted). Here, that review confirms the entire phrase should be limiting, including “the following steps applied at switching points distributed in said stream to decode a current symbol.”

As discussed at length above, what the '859 patent inventors allegedly invented was the concept of distributing “switching points” in the stream, so the encoder or decoder can “switch” to a probability model that is likely to be a better match for the source signal at that point in the stream. *See* Section III.B, *supra*. That was the entire basis of the claimed improvement over Arithmetic Coding, which uses the same probability model and updates it over time as the encoding or decoding process continues. *See id.*; JXM-2 ('859 Patent) at 2:23-33; RXM-41 (Orchard Decl.) at ¶¶ 38-46. According to the '859 patent, the approach of using of the same probability model over time was too slow to adapt to changing probabilities in the source signal itself. JXM-2 ('859 Patent) at 2:7-11 (describing that in Arithmetic Coding probabilities are

“regularly updated during the coding of the current signal in order to get closer to the real probability but with a delay time (AC case)”). The ’859 patent claims to provide an improvement by introducing the ability to switch to a *different* probability model at “switching points” distributed throughout the stream. *Id.* at 2:23-33. Accordingly, to the extent the phrase “the following steps applied at switching points distributed in said stream to decode a current symbol” is part of the preamble, it is limiting because it corresponds to “what the inventors actually invented and intended to encompass by the claim.” *Catalina Mktg.*, 289 F.3d at 808.

That “switching points” are at the heart of the alleged invention is also consistent with InterDigital’s own statements when prosecuting a related application in Europe. *See, e.g., Tanabe Seiyaku Co., Ltd. V. U.S. Int’l Trade Comm’n*, 109 F.3d 726, 733 (Fed. Cir. 1997) (finding that “representations made to foreign patent offices are relevant” to claim construction); *see also Starhome GmbH v. AT&T Mobility LLC*, 743 F.3d 849, 858 (Fed. Cir. 2014) (same). There, both the European Patent Office (“EPO”) and *InterDigital* understood it is “switching points” that provide the alleged benefits set forth in the ’859 patent specification.⁷ For example, the EPO stated:

The *invention* is defined so as to *switch the coding model* (which forms the basis for the data compression, for example the VLC data compression, like Huffman or arithmetic coding) *at so-called 'switching points'*.

RXM-7 (EP2449683 File History) at 157123. InterDigital never disputed that characterization of its alleged invention. Indeed, InterDigital later agreed that “both encoding and decoding methods

⁷ The patent application at issue there shared the same specification as the ’859 patent. RXM-7 (EP2449683 File History) at 157095–115.

have to know the locations of switching points in the stream” to carry out the invention. *Id.* at 157501.

Because selecting a probability model at “switching points” is “what the inventors actually invented and intended to encompass by the claim,” *Catalina Mktg.*, 289 F.3d at 808, the phrase “the following steps applied at switching points distributed in said stream to decode a current symbol” should be construed as limiting (to the extent it is deemed to be part of the preamble).

3. The Construction of “Switching Points”

The parties agree the specification includes a description of what “switching point” means in the context of the ’859 patent:

A switching point SP is defined according to the invention as a point in the stream at the level of which the probability model used to code the current symbol can be modified. The switching points SP are distributed in the stream, either uniformly or irregularly.

JXM-2 (’859 Patent) at 4:57-60 (emphasis added); *see also id.* at 9:56-59, 10:65-11:1. Both the private parties and Staff rely on this language for their proposed constructions.

The specification’s description of “switching points” does not resolve the dispute, however, because the parties disagree on what that description means. *See, e.g., Princeton Digital Image Corp. v. Konami Digital Entm’t Inc.*, 2017 WL 6375173, at *2 (D. Del. 2017) (adopting specification’s definition term and further construing the term “immersive” in the specification’s definition to resolve the parties’ dispute about what it means).⁸ In particular, the parties disagree

⁸ *See also Morpho Komodo LLC v. Blu Prods. Inc.*, 2016 WL 3976491, at *8 (E.D. Tex. 2016) (adopting specification’s definition of “signal” and further construing the definition); *Advanced Tech. Incubator, Inc. v. Sharp Corp.*, 2009 WL 749911, at *21 (E.D. Tex. 2009) (construing “passivating material” according to specification’s definition and further resolving disputes about the meaning of the definition); *Third Wave Techs., Inc. v. Eragen Bioscience, Inc.*, 2003 WL 23100277, at *5 (W.D. Wis. 2003) (recognizing that “[t]he definition [of the claim term at issue] provided in the specification does not resolve the dispute.”).

whether “switching points” are locations in the stream where a different probability model can be *selected*, or whether “switching points” also encompass locations in the stream where a probability model that is already in use can be *updated*. During meet and confers, Complainants have taken the position that their construction of “switching points” encompasses the latter—i.e., locations where a probability model is updated. Complainants’ construction should be rejected because it is contrary to the specification, it is contrary to representations that InterDigital itself made to foreign patent agencies, and it renders the entire concept of “switching points” meaningless. Respondents’ construction clarifies that “switching points” are located at points in the stream where “a probability model can be selected,” which is the entire point of the alleged invention.

As an initial matter, the claim language itself reflects that something (i.e., the probability model) is actually *switched* at the “switching point.” Respondents’ construction accounts for the claim language by reflecting that a probability model can be *selected* (e.g., switched out for a different one) at a “switching point.” See RXM-41 (Orchard Decl.) at ¶ 84. In contrast, Complainants’ construction would encompass points in the stream where nothing switched; the current probability model that is already in use is updated, but it is not switched.

The specification also supports Respondents’ construction. As discussed above, the ’859 patent repeatedly concedes that it was already known in the prior art to *update* a probability model as symbols are coded/decoded. See *id.* at 4:33-38 (describing that “AC entropy coding” uses the steps of coding a current symbol and then “updating the current probability model P_c according to the coding of the current symbol S_c .”); *id.* at 4:33-38. The patent makes clear that the entire purpose of the invention was to have “switching points” in the stream where the encoder/decoder can *select* a probability model that best matches the source signal at that point in time—not points

in the stream where the encoder/decoder merely *updates* a probability model that is already being used (as in the prior art). *See* Section III.B, *supra*; RXM-41 (Orchard Decl.) at ¶¶ 77-94.

According to the '859 patent, this prior art technique of *updating* the current probability model in use, i.e., as symbols are decoded over time, was too slow to adapt to fluctuating probabilities in the source signal. *See* JXM-2 ('859 Patent) at 2:7-11 (describing that in Arithmetic Coding probabilities are “regularly updated during the coding of the current signal in order to get closer to the real probability but with a delay time (AC case)” as “the real probability fluctuates during the coding process”). The '859 patent provides a purported improvement by introducing “switching points” into the stream that allow the encoder/decoder to “switch” out the probability model by selecting a different one, rather than just slowly adapting the current probability model to the source signal by updating it over time. *Id.* at 2:23-33. The patent explains that by doing this, the “convergence time” for the probability model to match the real probability of the source signal can be “considerably reduced with respect to the AC method.” *Id.* at 9:51-52, *see also* Figs. 2 and 9 (comparing convergence time of Arithmetic Coding (Fig. 2) with the alleged invention that uses “switching points” (Fig. 9)).

As shown in the below excerpts from the '859 patent specification, the critical distinction between the prior art (Arithmetic Coding) and the alleged invention is the use of “switching points” where a probability model can be selected. The '859 patent acknowledges that *both* Arithmetic Coding and the alleged invention use the same steps of **coding a current symbol** and **updating the current probability model**. What the alleged invention adds is “**switching points**” in the stream so the encoder/decoder can **select** a probability model that better matches the source signal:

'859 Patent's Description of Arithmetic Coding	'859 Patent's Description of Alleged Invention
“AC entropy coding comprises principally the	“ The invention relates to an arithmetic method for coding symbols in a stream, comprising the following steps:

following steps to code a current symbol Sc: coding of the current symbol Sc according to a current probability model Pc, and updating the current probability model Pc according to the coding of the current symbol Sc.” <i>Id.</i> at 4:33-38 (emphasis added).	coding a current symbol according to a current probability model, and updating the current probability model according to the coding of the current symbol. The method also comprises, <i>in the switching points distributed in the stream</i>, the following steps: selecting the current probability model in a set of at least two probability models according to a coding cost criterion, and coding an identifier of a selected probability model. Advantageously the arithmetic coding method according to the invention <i>enables the real probability to be approached more rapidly than with a standard VLC or AC type method</i> and thus enables the symbol or symbols to be coded more efficiently, i.e. at a lower coding cost.” <i>Id.</i> at 2:17-33 (emphasis added).
--	--

In this passage (above at right), the '859 patent again makes clear that “switching points” are locations where a probability model is *selected*. For example, after first describing that the alleged invention **codes** a symbol and **updates** the probability model just like Arithmetic Coding, the patent describes that the alleged invention “also” includes **selecting** a probability model at “**switching points** distributed in the stream,” which is what “enables the real probability to be approached more rapidly than with” Arithmetic Coding. *Id.* at 2:23-33; RXM-41 (Orchard Decl.) at ¶¶ 80-81. The passage further underscores that selecting a probability model at “switching points” is not just one embodiment; it is a feature of “**the invention**” itself (*id.* at 2:17, 2:30). *See, e.g., Luminara Worldwide, LLC v. Liown Elecs. Co. Ltd.*, 814 F.3d 1343, 1353 (Fed. Cir. 2016) (“When a patentee describes the features of the ‘present invention’ as a whole, he implicitly alerts the reader that this description limits the scope of the invention.” (internal quotation and citation omitted)). Indeed, the '859 patent describes *no embodiments* that do not involve selecting a probability model at a switching point.

The '859 patent's figures reinforce the same distinction between just updating a probability model after decoding each symbol (as in the prior art) and using “switching points” to *select* a probability model. As an example, Figure 11 shows a series of decoding steps that correspond to asserted claims 10 and 15.⁹

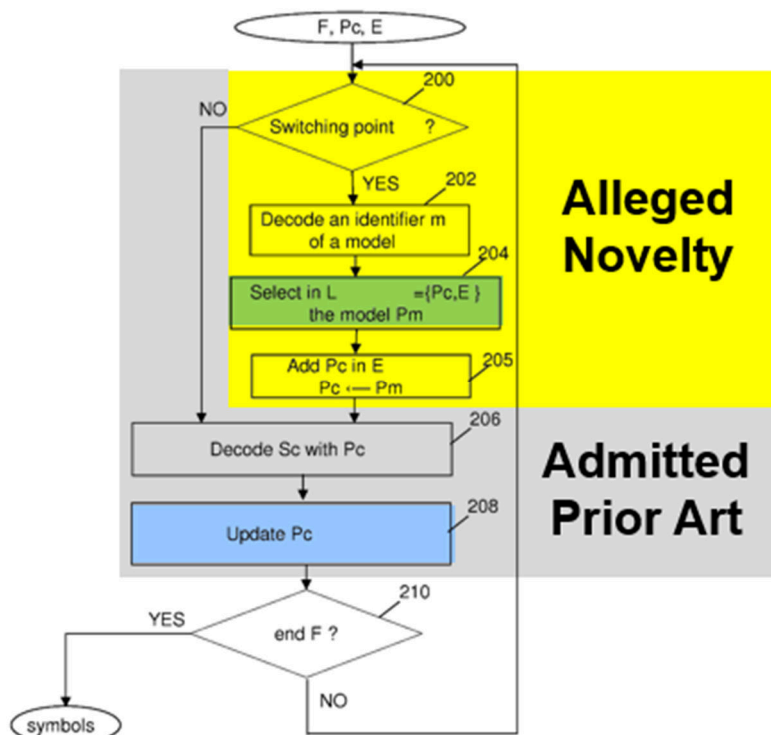


FIGURE 11

JXM-2 ('859 Patent) at Fig. 11 (annotated). The decoder first determines “if the location is on a switching point” (step 200) '859 patent at 10:64-65.

If the current symbol is *not located at a switching point* (the path highlighted in gray), it decodes the current symbol (step 206) and updates the current probability model (step 208). *Id.* at

⁹ Figures 10 and 11 of the '859 patent show two different embodiments for decoding a stream according to the invention. *See id.* at 4:21-24. Figure 11 shows the embodiment that corresponds to asserted claims 10 and 15, because it uses the claimed step of “adding” the current probability model. *See id.* at Fig. 11 (step 205), claim 10 (“adding . . .”) claim 15 (“means for adding . . .”). The embodiment shown in Figure 10 does not use the claimed “adding” step. *See id.* at Fig. 10.

11:55-59. Again, the '859 patent readily concedes those steps shown in Figure 11 are just borrowed from the prior art. *See id.* at 5:44-46 (describing that in Figure 3 “the current probability model P_c is updated according to the method described previously in reference to the AC coding.”); *see also id.* at 6:44-46, 8:55-57, 10:21-23, 11:57-59. On the other hand, if the current symbol is located at a switching point (the path highlighted in yellow), the decoder performs a series of steps to *select* a probability model for use going forward (steps 202, 204, and 205) before it then decodes the current symbol and updates the probability model using the steps known in the prior art (steps 206 and 208). *Id.* at 11:12-59.

As Figure 11 makes clear, “switching points” are *not* locations in the stream where a probability model can be updated (as Complainants contend). The probability model is *updated whether or not* the location is at a switching point. The difference in the two paths is clear: when the current symbol is at a “switching point,” the decoder actually *selects* a probability model to use. Thus, when the specification states that a “switching point” is “a point in the stream at the level of which the probability model used to code the current symbol can be modified,” JXM-2 ('859 Patent) at 4:57-60, it is referring to points in the stream where a different probability model can be selected to use in place of the current one, not just points in the stream where the current probability model can be updated. Indeed, Figure 11 shows the current probability model is always updated at *every location* in the stream, regardless of whether the current location is at a switching point. Complainants’ broad construction would render the entire concept of “switching points” meaningless, because every location in the stream would become a “switching point.”

Respondents’ construction plainly reflects what “switching points” actually are in the context of the invention, as *InterDigital itself has admitted* outside the context of this litigation. During the prosecution of the same related application discussed above, the EPO issued a rejection

because the specification does not disclose where the specifically the “switching points” should be located to achieve the invention’s alleged benefits. *See, e.g., Tanabe*, 109 F.3d at 733 (finding that “representations made to foreign patent offices are relevant” to claim construction); *see also Starhome*, 743 F.3d at 858 (same). In response, InterDigital pointed to the same specification language at issue here and explained that “switching points” are locations where the probability model can be modified “according to *selection* criterion”—i.e., where a probability model can be selected:

The definition of a switching point being “a point in the stream at the level of which the probability model used to code the current symbol *can be modified*” should be understand by a point in the stream at the level of which the probability model used to code the current symbol *is modified or not according to selection criterion*.

RXM-7 (EP2449683 File History) at 157276 (emphases added). InterDigital explained doubled down on this in yet another filing, confirming that “switching point” refers to a candidate point for “switching” the probability model and is what provides the alleged improvement over prior art:

This portion of the application specifies that the switching point is as a point in the stream at the level of which the probability model used to code the current symbol can be modified. In other words, *a switching point is a candidate point for switching the [probability model]* of a given symbol in the arithmetic coding/decoding. As previously submitted, the potential switching of [a probability model] at certain points, even if these are not exactly where the statistics change, allows a better coding efficiency, as the context in use up until the change in statistics, could take a longer time to adapt to such change, whereas the new context could guarantee an immediate fit.

Id. at 157500 (emphasis added).¹⁰ The EPO had the same understanding, stating that “switching points” are “points in the data stream” where “the probability models are *switched*, i.e., newly *selected*.” *Id.* at 157416 (emphasis added).¹¹ And later in prosecution, InterDigital again explained that switching points are located “where such new *selection* of probability models is to happen either in the encoding or decoding method.” *Id.* at 157502 (emphasis added). InterDigital said beyond that description, “switching points cannot otherwise be defined more precisely.” *Id.*

Finally, it is also worth noting that InterDigital adopted Respondents’ position when attempting to distinguish prior art. In particular, the EPO had rejected claims over “Context-Adaptive Binary Arithmetic Coding” (CABAC), which was a known prior art technique for arithmetic coding. InterDigital responded that CABAC merely disclosed a probability model is “continuously updated” during the encoding/decoding process. *Id.* at 157505. InterDigital claimed CABAC’s disclosure of continuously *updating* the probability model in use did not suggest “a possible *switching* of probability models at switching point in the stream.” *Id.* InterDigital explained that merely updating a probability model was not its invention, because updating occurs “not only at switching points but continuously with the coding of symbols.” *Id.* at 157506. In contrast, “switching points” are locations where a probability model can be switched, i.e., a different probability model can be selected.

¹⁰ The original language InterDigital used in this excerpt refers to a “context.” However, immediately before this excerpt, InterDigital explained the terms “context” and “probability model” have the same meaning and are interchangeable. *Id.* at 157499 (“A probability model is also known as context in the domain of arithmetic coding/decoding (AC).”).

¹¹ The EPO maintained its rejection because the specification did not disclose specifically where in the stream to put the “switching points.”

For the foregoing reasons, the term “switching point” should be construed as “a point in the stream at the level of which the probability model used to code the current symbol can be modified—i.e., a probability model can be selected.” InterDigital’s proposed construction is inconsistent with the specification, InterDigital’s own representations to foreign patent agencies, and the entire purpose of the invention.

B. “means for” terms of claim 15

Claim 15 of the ’859 patent recites five “means for” limitations. The parties agree these limitations are means-plus-function terms subject to 35 U.S.C. § 112 ¶ 6. The only disputed issues are whether the specification discloses sufficient structure for perform the claimed function, which is necessary to render the claim definite. It does not.

1. “means for decoding from the stream, at switching points distributed in said stream to decode a current symbol, for said current symbol, a probability model identifier” (claim 15)

Complainants’ Construction	Respondents’ Construction	Staff’s Construction
Subject to § 112, ¶ 6 Function: decoding from the stream . . . a probability model identifier Structure: an arithmetic decoder and equivalents thereof. See FIG. 13 and 13:25-52. Alternatively, the structure is an arithmetic decoder operable to perform the claimed function, as described in the specification in	Subject to § 112, ¶ 6. Function: decoding a probability model identifier for the current symbol from the stream, at switching points distributed in the stream Structure: indefinite for lack of sufficient corresponding structure.	Subject to § 112, ¶ 6 Function: at switching points in the stream, decoding a probability model identifier for the current symbol Structure: Fig. 13 and col. 13, ll. 25-52 disclose the structure of a decoding device and, more specifically, an entropy decoding module that implements this step/function. However, it is unclear whether this disclosure alone provides a sufficient structure to a person of ordinary skill in the art, such that they would know how to perform this function. Complainants cite to Figs. 10, 11, and 13 and 10:1-14, 11:12-34, and 13:25-52 as disclosing an arithmetic decoder operable to perform the claimed function in connection with step 202 and as describing step 202. These figures and cited paragraphs do not appear to disclose an algorithm that shows how the claimed function is

connection with step 202 and equivalents thereof, step 202 being described in the specification, for example at the following figures and passages: FIG. 13 and 13:25-52 and FIGS. 10, 11 and 10:1-14, 11:12-34.		carried out. However, absent expert testimony, it is not clear whether an algorithm must be disclosed in addition to the decoding device. In view of this, the Staff is not in a position to say that there is clear and convincing evidence showing indefiniteness. Further discovery, including expert discovery, may reveal that the patent does not disclose a sufficient corresponding structure, or it may narrow the precise portions of the specification that are the required structure.
--	--	--

For this term, the parties dispute both the claimed function and whether the specification discloses sufficient structure for performing the function.

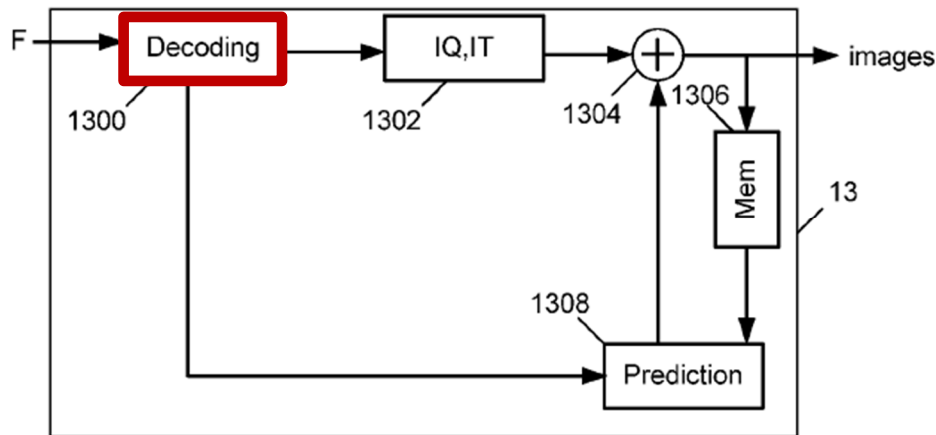
As an initial matter, Complainants truncate the claimed function by shortening it to just “decoding from the stream a probability model identifier,” whereas the claim language recites the additional requirements that the action of decoding a probability model identifier is performed “for the current symbol” and “at switching points distributed in the stream.” These limitations provide additional details about the action taken (when it is taken, the result of taking it, what it operates on, etc.), are part of the claimed function, and should be identified as such in the means-plus-function analysis. *See, e.g., Bennett Marine, Inc. v. Lenco Marine, Inc.*, 549 Fed. Appx. 947, 954 (Fed. Cir. 2013) (finding the “district court erred by failing to fully identify the function” because the claim recited that the “means” must “automatically mov[e]” the trim tabs “upon removal of power at said engine”) (citing *Lockheed Martin Corp. v. Space Sys./Loral, Inc.*, 324 F.3d 1308, 1319 (Fed. Cir. 2003) (“The function of a means-plus-function claim must be construed to include the limitations contained in the claim language.”)); *EMA Electromechanics, Inc. v. Siemens Corp.*, No. 6:21-CV-1001-ADA, 2022 WL 526244, at *6 (W.D. Tex. Feb. 16, 2022) (finding that plaintiff’s “truncated” function where “does not accurately reflect the functional requirements” recited in the claim because “[r]ather than describing general movement of the actuator arm, the

function recited in [the claim] describes moving the actuator arm between two distinct positions resulting in specific connections for certain contactors.”). Respondents’ and Staff’s proposed functions are similar and account for these additional limitations that specify what the claimed function is, while Complainants’ proposed function reads them out.

Regardless of which party’s proposed function is adopted, the specification fails to recite sufficient structure under § 112 ¶ 6. That is, the specification fails to disclose sufficient structure for performing the more general function of “decoding from the stream a probability model identifier” (which Complainants propose), much less for performing the more specific function of doing so “for the current symbol” and “at switching points distributed in the stream” (as reflected in Respondents’ and Staff’s proposals). Although the specification discusses “decoding” a probability model identifier from the stream, it does so by describing the function and not by describing any particular structure for performing it. *See* JXM-2 (’859 Patent) at Figs. 10 & 11, 10:1-14, 11:12-34.

Complainants rely on two “alternative” sets of disclosures to argue the specification discloses sufficient structure corresponding to the “decoding” function. Respondents address each set of disclosures below.

First, Complainants identify the corresponding structure as “an arithmetic decoder and equivalents thereof. *See* FIG. 13 and 13:25-52.” Those parts of the ’859 patent do not recite sufficient corresponding structure, however. RXM-41 (Orchard Decl.) at ¶¶ 101-09. According to the patent, the outermost box shown in Figure 13 is a “decoding device” (labeled 13). JXM-2 (’859 Patent) at 4:27-28. Complainants apparently contend this “decoding device” and its components (labeled 1300, 1302, 1304, 1306, and 1308) show a structure corresponding to the “decoding” function at issue in this term:



Id. at Figure 13 (annotated). Complainants are incorrect. Figure 13 does not provide any information about what structure(s) correspond to the claimed “decoding” function. As explained below, the ’859 patent states the “decoding” function is performed within the “**entropy decoding module**” labeled 1300.

Indeed, the patent explains that the functions recited in claim 15 (not only for this means-plus-function term, but for all of the others addressed below) are all performed within **entropy decoding module** 1300. To understand why, it is helpful to refer back to Figure 11, shown at right. As discussed above, Figure 11 shows the series of decoding steps recited in claim 15’s means-plus-function limitations. For example, Figure 11 shows a series of steps performed at a switching point, including decoding a probability model identifier (step 202), selecting a probability model using the identifier (step 204), adding a probability model into the subset of probability models (step 205), decoding the current symbol with the current

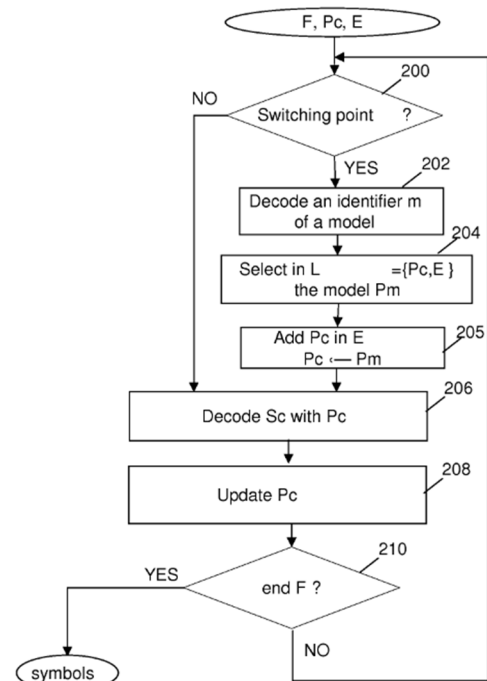


FIGURE 11

probability model (step 206), and updating the current probability model (step 208). When later discussing Figure 13, the '859 patent states that these steps shown in Figure 11 are all performed within the **entropy decoding module** 1300 shown in Figure 13. *Id.* at 13:32-35 (“[T]he entropy decoding module 1300 implements the steps 200, 202, 204, 206, 208, 210 and possibly 205 of the arithmetic decoding method according to the invention.”).¹²

In sum, the “decoding device” 13 shown in Figure 13 is not where the claimed functions are performed; the claimed functions are performed within “entropy decoding module” 1300. Yet, Figure 13 shows nothing about the structure of entropy decoding module 1300; it is a complete black box labeled “decoding,” which just restates the claimed function itself. Accordingly Figure 13 is insufficient to provide the required corresponding structure. *See Noah Sys.*, 675 F.3d at 1316–17 (“The portions of the specification that describe what occurs inside box 44, however, merely recite functional, not structural, language. . . . This type of purely functional language, which simply restates the function associated with the means-plus-function limitation, is insufficient to provide the required corresponding structure.”).

Nor do the words “entropy decoding module” themselves inform a POSITA what specific structures correspond to the claimed functions. As an initial matter, “module” is a “well-known nonce word” that does not disclose any structure, just like “means.” *Williamson*, 792 F.3d at 1350 (finding that “[m]odule” is a well-known nonce word that can operate as a substitute for ‘means’ in the context of § 112, para. 6” and that it “is simply a generic description for software or hardware that performs a specified function.”); *Ergo Licensing, LLC v. CareFusion 303, Inc.*, 673 F.3d 1361, 1363–64 (Fed. Cir. 2012) (rejecting argument that the specification’s recitation

¹² Although the '859 patent states that step 205 is “possibly” implemented in entropy decoding module 1300, the patent does not describe any other location where it would be implemented.

of a “control device” disclosed sufficient structure for the claimed “control means” because “[t]he recitation of ‘control device’ provides no more structure than the term “control means” itself, rather it merely replaces the word “means” with the generic term “device.”).

Even setting aside the word “module” and focusing on the rest of the phrase, the words “entropy decoding” also do not render the term definite. A POSITA would understand that decoding, including entropy or arithmetic decoding, can be implemented in any number of different structures, including a microprocessor, discrete logic, an application specific integrated circuit (ASIC), a field programmable gate array (FPGA), a digital signal processor (DSP), a system on a chip (SoC), or discrete logic, software, hardware, firmware or any combinations thereof. RXM-41 (Orchard Decl.) at ¶¶ 104-06. In the below table, Respondents provide just a few examples of extrinsic evidence confirming that a POSITA would have understood there to be many possible structures for implementing a decoder:

Exhibit	U.S. Patent	Title	Options for Implementing Decoder
RXM-1	7,932,843	Parallel CABAC decoding for video decompression	“Any suitable combination of software, firmware, and hardware” such as, one or more digital signal processors (DSPs), microprocessors, discrete logic, application specific integrated circuits (ASICs), systems on a chip (SoC). <i>See, e.g.</i> , 5:12-17, 16:4-25.
RXM-2	8,351,502	Method and apparatus for adaptively selecting context model for entropy coding	General purpose computer, special purpose computer, Field Programmable Gate Array (FPGA), ASIC. <i>See, e.g.</i> , 6:9-17, 16:12-32.
RXM-3	8,416,857	Parallel or pipelined macroblock processing	FPGA, DSP, general purpose digital computer, ASIC, or programmable logic devices. <i>See, e.g.</i> , 4:17-24, 12:47-49, 13:46-55.
RXM-4	8,718,149	Image coding method and image decoding method	A plurality of processors or in a single processor, a Large Scale Integration (LSI), FPGA. <i>See, e.g.</i> , 41:16-21, 53:6-24.

RXM-5	9,008,171	Two pass quantization for CABAC coders	Microprocessor, ASIC, FPGA, DSP. <i>See, e.g.</i> , 3:26-34, 6:24-29, 20:63-21:3.
RXM-6	9,386,316	Adaptive coding of video block header information	Microprocessor, ASIC, FPGA, DSP, “or other equivalent integrated or discrete logic circuitry.” <i>See, e.g.</i> , 3:23-33, 9:13-17, 25:14-21.

The ’859 patent itself admits the same: there is no particular, physical structure that a POSITA would recognize as corresponding to the claimed functions. According to the ’859 patent, the claimed functions are “are not linked to a coding standard,” *id.* at 13:67-14:2, and they could be implemented in any number of physical forms, including “in a single component,” as “functions of the same software,” or as “separate physical entities,” *id.* at 13:53-61. The patent says nothing more about what those structures would actually look like, however. Under these circumstances, the mere reference to “entropy decoding” does not disclose sufficient structure to render the claim definite.

The Federal Circuit’s decision in *Ergo Licensing, LLC v. CareFusion 303, Inc.*, 673 F.3d 1361 (Fed. Cir. 2012), is instructive. There, the court held the specification’s description of a “control device” insufficient to provide structural support for the claimed “control means” because “there were *at least three different types of control devices commonly available and used* at the time to control adjusting means: microprocessors, discrete circuits connected to stepper motors, and analog circuits.” *Id.* at 1364 (emphasis added). The court explained that “[a]lthough one of skill in the art may have been able to find a structure that would work, that does not satisfy § 112 ¶ 6. Under § 112 ¶ 6, a patentee is only entitled to ‘corresponding structure . . . described in the specification and equivalents thereof,’ not any device capable of performing the function.” *Id.* (quoting 35 U.S.C. § 112 ¶ 6 (emphasis in original)); *see also Synchronoss Techs., Inc. v. Dropbox, Inc.*, 987 F.3d 1358, 1367–68 (Fed. Cir. 2021) (finding “user identifier module”

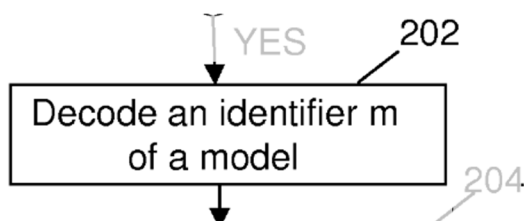
indefinite where the specification did not disclose any corresponding structure and a POSITA would understand there to be “many ways in which a system could perform user identification, each with its own distinct structure”). The Federal Circuit has explained that it is “not enough that a means-plus-function claim term correspond to every known way of achieving the claimed function; instead, the term must correspond to ‘adequate’ structure *in the specification* that a person of ordinary skill in the art would be able to recognize and associate with the corresponding function in the claim.” *Synchronoss*, 987 F.3d at 1367.

As set forth in *Ergo*, the relevant question here is not whether a POSITA “may have been able to find a structure that would work.” *Ergo*, 673 F.3d at 1364. The question is whether the *specification itself* discloses a structure that a POSITA would be able to recognize and associate with the function in the claim. *Id.*; *Synchronoss*, 987 F.3d at 1367. Where the specification uses generic descriptions that could be implemented in any number of different structures, it fails to provide the guidance that § 112 ¶ 6 requires. *Ergo*, 673 F.3d at 1364. Because a POSITA would have understood that a number of different structures could be used to implement entropy decoding (or arithmetic decoding), but the specification actually *describes* none of them, the ’859 patent fails to recite sufficient corresponding structure to render this term definite.

Finally, the specification passage that Complainants cite (13:25-52) also just describes decoding in functional terms. The passage merely states that the entropy decoding module is “able to generate decoded data, for example coding modes and decoded data relating to the content of the images.” Simply restating the *function* of the entropy decoding module, i.e., decoding, does not inform a POSITA what *structures* correspond to the claimed function. *Finisar Corp. v. DirecTV Group, Inc.*, 523 F.3d 1323, 1340 (Fed. Cir. 2008) (finding no

structural disclosure in a passage that “provides ‘nothing more than a restatement of the function, as recited in the claim.’”)

Second, in the alternative, Complainants assert the corresponding structure is “an arithmetic decoder operable to perform the claimed function, as described in the specification in connection with step 202 and equivalents thereof, step 202 being described in the specification, for example at the following figures and passages: FIG. 13 and 13:25-52 and FIGS. 10, 11 and 10:1-14, 11:12-34.” These alternative disclosures also do not disclose sufficient structure. RXM-41 (Orchard Decl.) at ¶¶ 109-12. The “step 202” that Complainants identify (shown at right in Figure 11) is simply a black box that restates the *function* of “decoding” a probability model identifier.



Indeed, the '859 patent readily admits that step 202 shows a “functional unit.” *Id.* at 13:53-55. Merely restating the claimed function does not render the claim definite, as a matter of law. *See Noah Sys.*, 675 F.3d at 1316–17 (“The portions of the specification that describe what occurs inside box 44, however, merely recite functional, not structural, language. . . . This type of purely functional language, which simply restates the function associated with the means-plus-function limitation, is insufficient to provide the required corresponding structure.”); *Finisar*, 523 F.3d at 1340 (finding no structural disclosure in a passage that “provides nothing more than a restatement of the function, as recited in the claim.” (internal quotation omitted)).

Nor do the passages Complainants cite (10:1-14 and 11:12-34) disclose any structure. Both passages merely state that during step 202, “an identifier m is decoded from the stream” and that the identifier “identifier enables identification of a probability model in a set L of at least two probability models.” JXM-2 ('859 Patent) at 10:1-14, 11:12-34. The remainder of the passages

just discuss what is included in the set of probability models. *See id.* Here again, this mere restatement of the claimed *function* (“decoding” a probability model identifier) does not inform a POSITA of corresponding *structures* are used for decoding and does not render the claim definite.

* * *

For the foregoing reasons, neither of Complainants’ alternative sets of disclosures provides sufficient structure corresponding to the “decoding” function. Complainants cannot fix this deficiency with expert gloss, since expert testimony cannot create structure where none is adequately disclosed in the specification itself, which is what § 112 ¶ 6 requires. *See Williamson*, 792 F.3d at 1354 (“The testimony of one of ordinary skill in the art cannot supplant the total absence of structure from the specification. . . . The prohibition against using expert testimony to create structure where none otherwise exists is a direct consequence of the requirement that the specification adequately disclose corresponding structure.”). The specification’s failure to disclose sufficient corresponding structure for this term renders claim 15 indefinite.

2. **“means for selecting using said probability model identifier a probability model in a set comprising a current probability model defined from symbols coded previous to the current symbol and a subset of probability models wherein subset comprises at least one probability model defined by default” (claim 15)**

Complainants’ Construction	Respondents’ Construction	Staff’s Construction
Subject to § 112, ¶ 6 Function: selecting . . . a probability model Structure: an arithmetic decoder and equivalents thereof. See FIG. 13 and 13:25-52.	Subject to § 112, ¶ 6 Function: using the probability model identifier to select a probability model in a set comprising a current probability	Subject to § 112, ¶ 6 Function: selecting a probability model using the probability model identifier, where the probability model is selected from a current probability model and a subset of probability models, including a default probability model Structure: Fig. 13 and col. 13, ll. 25-52 disclose the structure of a decoding device

Alternatively, the structure is an arithmetic decoder operable to perform the claimed function, as described in the specification in connection with step 204 and equivalents thereof, step 204 being described in the specification, for example at the following figures and passages: FIG. 13 and 13:25-52 and FIGS. 10, 11 and 5:36-43, 10:1-17, 11:12-36.	model defined from symbols coded previous to the current symbol and a subset of probability models wherein said subset comprises at least one probability model defined by default Structure: indefinite for lack of sufficient corresponding structure.	and, more specifically, an entropy decoding module that implements this step/function. Complainants cite to Figs. 10, 11, and 13 and 5:36-43, 10:1-17, 11:12-36, and 13:25-52 as disclosing an arithmetic decoder operable to perform the claimed function in connection with step 204 and as describing step 204. It is unclear whether these figures and cited paragraphs disclose an algorithm that shows how the claimed function is carried out. However, absent expert testimony, it is not clear whether an algorithm must be disclosed in addition to the decoding device. Further discovery, including expert discovery, may reveal that the foregoing disclosures provide a sufficient structure, or may reveal that the patent does not disclose a sufficient corresponding structure.
--	---	---

The issues for this claim term are similar to the “means for decoding . . . a probability model identifier” term discussed above.

As an initial matter, Complainants truncate the claimed function by shortening it to “selecting a probability model,” whereas the claim language recites the additional requirements that the action of selecting a probability model is performed “using said probability model identifier” and that the selection is further to select a probability model “in a set comprising a current probability model defined from symbols coded previous to the current symbol and a subset of probability models wherein subset comprises at least one probability model defined by default.” For the same reasons described above, these additional details about the action taken are part of the claimed function and should be identified as such in the means-plus-function analysis. *See Bennett Marine*, 549 Fed. Appx. at 954 (Fed. Cir. 2013) (finding the “district court erred by failing to fully identify the function” because the claim recited that the “means” must “automatically mov[e]” the trim tabs “upon removal of power at said engine”). While Respondents’ and Staff’s proposed functions are similar, Respondents respectfully submit that their proposed function

accounts for the entire phrase at issue, whereas Staff's proposed function omits certain aspects (e.g., that the current probability model that is part of the selection is "defined from symbols coded previous to the current symbol").

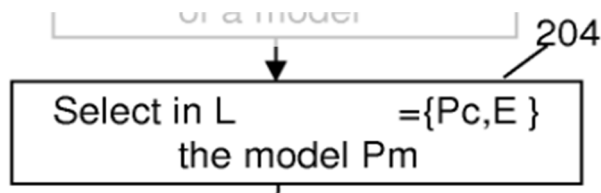
Nevertheless, as with the preceding term, the specification fails to recite sufficient structure regardless of which party's proposed function is adopted. The specification fails to disclose sufficient structure for performing the more general function of "selecting a probability model" that Complainants propose, much less for performing the more specific functions reflected in Respondents' and Staff's proposals. Although the specification discusses "selecting" a probability model, it does so by describing the function and not by describing any particular structure for performing it. *See* JXM-2 ('859 Patent) at Figs. 10 & 11, 10:1-17, 11:12-36.

Complainants rely on two "alternative" sets of disclosures to argue the specification discloses sufficient structure corresponding to the "selecting" function. Respondents address each set of disclosures below.

First, Complainants identify the corresponding structure as "an arithmetic decoder and equivalents thereof. *See* FIG. 13 and 13:25-52," which is the same structure Complainants identified for the "means for decoding . . . a probability model identifier" term above. For the reasons discussed above, those portions of the '859 patent do not recite sufficient corresponding structure. The specification's description of the "selecting" step is entirely functional. The fact that the specification states the "selecting" occurs within a black box called an "entropy decoding module" 1300 does not sufficiently inform a POSITA of any corresponding structure for the same reasons discussed above. *See* Section VI.B.1; RXM-41 (Orchard Decl.) at ¶ 116.

Second, in the alternative, Complainants assert the corresponding structure is “an arithmetic decoder operable to perform the claimed function, as described in the specification in connection with step 204 and equivalents thereof, step 204 being described in the specification, for example at the following figures and passages: FIG. 13 and 13:25-52 and FIGS. 10, 11 and 5:36-43, 10:1-17, 11:12-36.” Again, these alternative disclosures also do not disclose sufficient structure. RXM-41 (Orchard Decl.) at ¶¶ 117-

21. The “step 204” that Complainants identify (shown at right in Figure 11) is simply a black



box that restates the *function* of “selecting” a probability model. The ’859 patent admits that step 204 shows a “functional unit.” JXM-2 (’859 Patent) at 13:53-55.

Nor do the passages Complainants cite (10:1-17 and 11:12-36) disclose any structure. Both passages merely restate the *function* that during step 204, “the probability model of L for which the identifier is m is selected.” *Id.* at 10:1-17 and 11:12-36. Finally, Complainants cite one additional passage for this term (5:36-43), but that passage is also unavailing. Like the two other passages addressed above, it merely describes a *function*—namely, the function that during step 104 of Figure 3, “an identifier m of the model selected in the set L is coded in the stream.” *Id.* at 5:36-37. In addition, that passage is not relevant to the function actually at issue for this term, which relates to *selecting* a probability model using a probability model identifier. The passage at 5:36-43 instead describes the function of *coding* an identifier into the stream for the selected probability model. *Id.* at 5:36-43.

3. “means for adding in said subset said current probability model, the selected probability model becoming the current probability model” (claim 15)

Complainants’ Construction	Respondents’ Construction	Staff’s Construction
Subject to § 112, ¶ 6 Function: adding in said subset said current probability model Structure: an arithmetic decoder and equivalents thereof. See FIG. 13 and 13:25-52. Alternatively, the structure is an arithmetic decoder operable to perform the claimed function, as described in the specification in connection with step 205 and equivalents thereof, step 205 being described in the specification, for example at the following figures and passages: FIG. 13 and 13:25-52 and FIGS. 8, 10, 11 and 11:37-43 (describing 205 and additionally that step 205 can be identical to step 105), FIG. 7 and 8:8-37 (describing step 105).	Subject to § 112, ¶ 6 Function: adding the current probability model into the subset of probability models Structure: indefinite for lack of sufficient corresponding structure.	Subject to § 112, ¶ 6 Function: adding the current probability model into the subset of probability models Structure: Fig. 13 and col. 13, ll. 25-52 disclose the structure of a decoding device and, more specifically, an entropy decoding module that possibly implements this step/function. Complainants cite to Figs. 7, 8, 10, 11, and 13 and 11:37-43 (describing 205 and additionally that step 205 can be identical to step 105) and 8:8-37 (describing step 105) as disclosing an arithmetic decoder operable to perform the claimed function in connection with step 205 and as describing step 205. Fig. 8 and col. 11, ll. 37-43 appear to disclose an algorithm for performing this step. However, absent expert discovery, it is unclear whether this disclosure provides a sufficient structure to a person of ordinary skill in the art, such that they would know how to perform this function.

The issues for this claim term are similar to the “means for decoding . . . a probability model identifier” term discussed above. Each party’s proposed function is similar in substance, providing that the function is adding the current probability model into the subset of probability models.

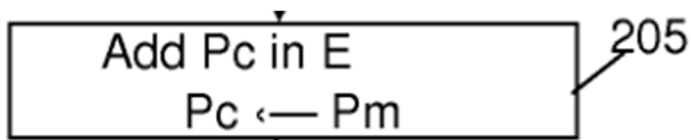
As with the preceding term, the specification fails to recite sufficient structure regardless of which party’s proposed function is adopted. The specification fails to disclose sufficient structure for performing the more general function of “selecting a probability model” that

Complainants propose, much less for performing the more specific functions reflected in Respondents’ and Staff’s proposals. Although the specification discusses “selecting” a probability model, it does so by describing the function and not by describing any particular structure for performing it. *See* JXM-2 (’859 Patent) at Fig. 11, 11:37-43.

Complainants rely on two “alternative” sets of disclosures to argue the specification discloses sufficient structure corresponding to the “adding” function. Respondents address each set of disclosures below.

First, Complainants identify the corresponding structure as “an arithmetic decoder and equivalents thereof. *See* FIG. 13 and 13:25-52,” which is the same structure Complainants identified for the “means for decoding . . . a probability model identifier” term above. For the reasons discussed above, those portions of the ’859 patent do not recite sufficient corresponding structure. The specification’s description of the “adding” step is entirely functional. The fact that the specification states the “selecting” occurs within a black box called an “entropy decoding module” 1300 does not sufficiently inform a POSITA of any corresponding structure for the same reasons discussed above. *See* Section VI.B.1; RXM-41 (Orchard Decl.) at ¶ 125.

Second, in the alternative, Complainants assert the corresponding structure is “an arithmetic decoder operable to perform the claimed function, as described in the specification in connection with step 205 and equivalents thereof, step 205 being described in the specification, for example at the following figures and passages: FIG. 13 and 13:25-52 and FIGS. 8, 10, 11 and 11:37-43 (describing 205 and additionally that step 205 can be identical to step 105), FIG. 7 and 8:8-37 (describing step 105).” Again, these alternative disclosures also do not disclose sufficient structure. RXM-41 (Orchard Decl.) at ¶¶ 126-30. The “step 205” that Complainants



identify (shown at right in Figure 11) is simply a black box that restates the *function* of “adding” the current probability model. The ’859 patent admits that step 205 shows a “functional unit.” JXM-2 (’859 Patent) at 13:53-55.

Nor does the cited passage at 11:37-43 disclose any structure. There, the specification merely restates the *function* that during step 205, “the current probability model P_c is added to the subset E .” *Id.* at 11:37-43. Although the specification states that this step is “further detailed” in Figure 8, that figure does not identify any corresponding structure, either. To the extent Figure 8 describes “adding” the current probability model, it does so using a generic, black box description, as shown below in step 1054:

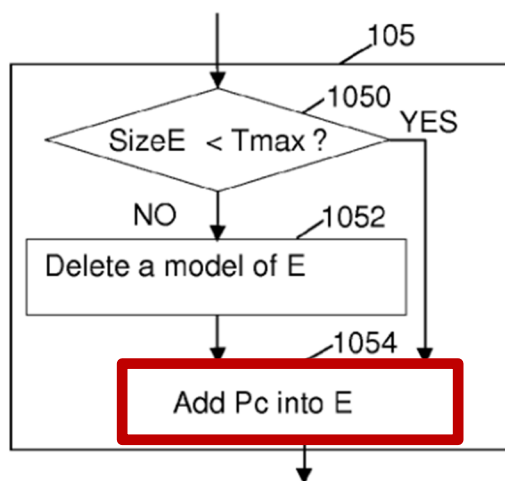
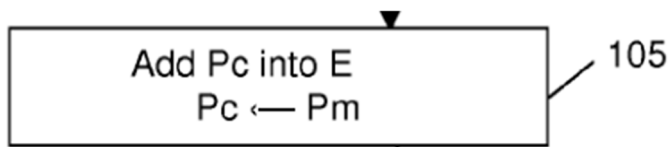


FIGURE 8

Id. at Figure 8 (annotated).

Nor do the remaining disclosures (Figure 7 and 8:8-37) help Complainants. Like Figures 8, 10, and 11, step 105 of Figure 7 (shown at right) just restates the function of the “adding” the current probability model. Similarly, the specification at 8:8-37 just describes the *function* that during step 105, “the current probability model P_c is added to the subset E .” *Id.* at



8:9-10. The passage then continues by describing some alleged benefits the “adding” function (e.g., that it makes the coding process “more efficient,” *id.* at 8:14-17), but it does not describe any *structure* corresponding to the “adding” function that would render this term definite.

4. “means for decoding said current symbol with said current probability model” (claim 15)

Complainants’ Construction	Respondents’ Construction	Staff’s Construction
Subject to § 112, ¶ 6 Function: decoding said current symbol Structure: an arithmetic decoder and equivalents thereof. See FIG. 13 and 13:25-52. Alternatively, the structure is an arithmetic decoder operable to perform the claimed function, as described in the specification in connection with step 206 and equivalents thereof, step 206 being described in the specification, for example at the following figures and passages: FIG. 13 and 13:25-52 and FIGS. 10, 11 and 10:18-19, 11:55-57.	Subject to § 112, ¶ 6 Function: decoding the current symbol with the current probability model Structure: indefinite for lack of sufficient corresponding structure.	Subject to § 112, ¶ 6 Function: decoding the current symbol with the current probability model Structure: Fig. 13 and col. 13, ll. 25-52 disclose the structure of a decoding device and, more specifically, an entropy decoding module that implements this step/function. However, it is unclear whether this disclosure alone provides a sufficient structure to a person of ordinary skill in the art, such that they would know how to perform this function. Complainants cite to Figs. 10, 11, and 13 and 10:18-19, 11:55-57, and 13:25-52 as disclosing an arithmetic decoder operable to perform the claimed function in connection with step 206 and as describing step 206. These figures and cited paragraphs do not appear to disclose an algorithm that shows how the claimed function is carried out, but absent expert testimony, the Staff is not in a position to say that there is clear and convincing evidence showing indefiniteness. Further discovery, including expert discovery, may reveal that the patent does not disclose a corresponding structure, or it may narrow the precise portions of the specification that are the required structure.

The issues for this claim term are similar to the “means for decoding . . . a probability model identifier” term discussed above.

As an initial matter, Complainants truncate the claimed function by shortening it to “decoding said current symbol,” whereas the claim language recites the additional requirements that the action of decoding the current symbol is performed “with the current probability model.” For the same reasons described above, this additional detail about the action taken is part of the claimed function and should be identified as such in the means-plus-function analysis. *See Bennett Marine*, 549 Fed. Appx. at 954 (Fed. Cir. 2013) (finding the “district court erred by failing to fully identify the function” because the claim recited that the “means” must “automatically mov[e]” the trim tabs “upon removal of power at said engine”). Respondents’ and Staff’s proposed functions are the same and account for the entire claimed function, while Complainants’ proposed function reads part of it out.

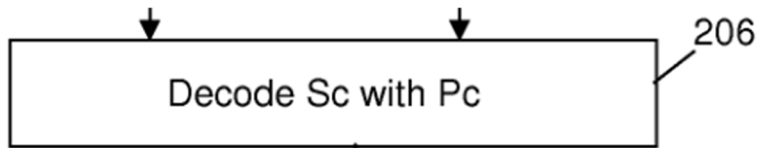
Regardless of which party’s proposed function is adopted, the specification fails to recite sufficient structure under § 112 ¶ 6. That is, the specification fails to disclose sufficient structure for performing the more general function of “decoding said current symbol” (which Complainants propose), much less for performing the more specific function of “decoding the current symbol with the current probability model” (as reflected in Respondents’ and Staff’s proposals). Although the specification discusses “decoding” current symbol, it does so by describing the function and not by describing any particular structure for performing it. *See JXM-2* (’859 Patent) at Figs. 10 & 11, 10:18-19, 11:55-57.

Complainants rely on two “alternative” sets of disclosures to argue the specification discloses sufficient structure corresponding to the “decoding” function. Respondents address each set of disclosures below.

First, Complainants identify the corresponding structure as “an arithmetic decoder and equivalents thereof. *See* FIG. 13 and 13:25-52,” which is the same structure Complainants

identified for the “means for decoding . . . a probability model identifier” term above. For the reasons discussed above, those portions of the ’859 patent do not recite sufficient corresponding structure. The specification’s description of the “decoding” step is entirely functional. The fact that the specification states the “selecting” occurs within a black box called an “entropy decoding module” 1300 does not sufficiently inform a POSITA of any corresponding structure for the same reasons discussed above. *See* Section VI.B.1; RXM-41 (Orchard Decl.) at ¶ 134.

Second, in the alternative, Complainants assert the corresponding structure is “an arithmetic decoder operable to perform the claimed function, as described in the specification in connection with step 206 and equivalents thereof, step 206 being described in the specification, for example at the following figures and passages: FIG. 13 and 13:25-52 and FIGS. 10, 11 and 10:18-19, 11:55-57.” Again, these alternative disclosures also do not disclose sufficient structure. RXM-41 (Orchard Decl.) at ¶¶ 135-38. The “step 206” that Complainants identify (shown at right in Figure 11) is simply a black box



that restates the *function* of “decoding” a current symbol with the current probability model. The ’859 patent admits that step 206 shows a “functional unit.” JXM-2 (’859 Patent) at 13:53-55. Nor do the two new passages Complainants cite (10:18-19 and 11:55-57) disclose any structure. Both passages merely restate the *function* that during step 206, “the current symbol Sc is decoded with the current probability model Pc.” *Id.* at 10:18-19; 11:55-57.

5. “means for updating said current probability model according to the decoding of said current symbol” (claim 15)

Complainants’ Construction	Respondents’ Construction	Staff’s Construction
Subject to § 112, ¶ 6	Subject to § 112, ¶ 6	Subject to § 112, ¶ 6

Function: updating said current probability model Structure: an arithmetic decoder and equivalents thereof. See FIG. Figure 13 and 13:25-52. Alternatively, the structure is an arithmetic decoder operable to perform the claimed function, as described in the specification in connection with step 208 and equivalents thereof, step 208 being described in the specification, for example at the following figures and passages: FIG. 13 and 13:25-52 and FIGS. 10, 11, 4:33-54, 10:20-22, 11:57-61.	Function: updating the current probability model according to the decoding of the current symbol Structure: an algorithm wherein an occurrence number for the current symbol is increased by 1	Function: updating the current probability model according to the decoding of the current symbol Structure: Fig. 13 and col. 13, ll. 25-52 disclose the structure of a decoding device and, more specifically, an entropy decoding module that implements this step/function. However, it is unclear whether this disclosure alone provides a sufficient structure to a person of ordinary skill in the art, such that they would know how to perform this function. Complainants cite to Figs. 10, 11, and 13 and 4:33-54, 10:20-22, 11:57, and 13:25-52 as disclosing an arithmetic decoder operable to perform the claimed function in connection with step 208 and as describing step 208. Col. 4, ll. 33-54 describes “AC entropy coding.” However, absent expert discovery, it is unclear whether this disclosure provides a sufficient structure to a person of ordinary skill in the art, such that they would know how to perform this function.
--	--	---

While the issues for this claim term are similar to the other means-plus-function terms discussed above, Complainants do not assert that this term is indefinite under § 112 ¶ 6, since the specification arguably discloses structure (in the form of an algorithm) for “updating said current probability model according to the decoding of said current symbol.”

As an initial matter, Complainants truncate the claimed function by shortening it to “updating said current probability model,” whereas the claim language recites the additional requirements that the action of updating the current probability model is performed “according to the decoding of the current symbol.” For the same reasons described above, this additional detail about the action taken is part of the claimed function and should be identified as such in the means-plus-function analysis. *See Bennett Marine*, 549 Fed. Appx. at 954 (Fed. Cir. 2013) (finding the “district court erred by failing to fully identify the function” because the claim recited that the

“means” must “automatically mov[e]” the trim tabs “upon removal of power at said engine”). Respondents’ and Staff’s proposed functions are the same and account for the entire claimed function, while Complainants’ proposed function reads part of it out.

When describing the claimed “updating” step, the ’859 patent specification states the current probability model is updated “according to the method described previously in reference to the AC coding.” *Id.* at 11:57-59 (describing step 208 in Figure 11). The “described previously” is a reference to earlier in the specification, where the patent describes that Arithmetic Coding updates a probability model by increasing the occurrence number for the current symbol by 1:

If the symbol S_c takes the value ‘b’ then *in the step of updating of the current probability model P_c , the associated probability is increased, i.e. the occurrence number $F(S_c=b)$ is increased by 1.*

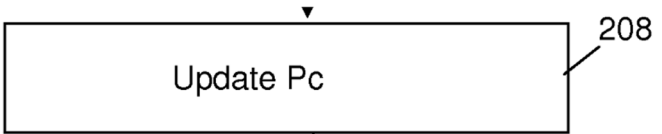
’859 patent, at 4:42–45. Since this is the only arguable disclosure of an “algorithm” for updating a probability model according to the decoding of the current symbol (i.e., the claimed function), the corresponding structure for this term should be construed as “an algorithm wherein an occurrence number for the current symbol is increased by 1.” *See* RXM-41 (Orchard Decl.) at ¶¶ 141-42.

As with the foregoing means-plus-function terms, Complainants propose two sets of disclosures as corresponding structure. Respondents address each set of disclosures below.

First, Complainants identify the corresponding structure as “an arithmetic decoder and equivalents thereof. *See* FIG. 13 and 13:25-52,” which is the same structure Complainants identified for the “means for decoding . . . a probability model identifier” term above. For the reasons discussed above, those portions of the ’859 patent do not recite sufficient corresponding structure. The fact that the specification states the “selecting” occurs within a black box called an “entropy decoding module” 1300 does not sufficiently inform a POSITA of any corresponding

structure for the same reasons discussed above. *See* Section VI.B.1; RXM-41 (Orchard Decl.) at ¶ 144.

Second, in the alternative, Complainants assert the corresponding structure is “an arithmetic decoder operable to perform the claimed function, as described in the specification in connection with step 208 and equivalents thereof, step 208 being described in the specification, for example at the following figures and passages: FIG. 13 and 13:25-52 and FIGS. 10, 11, 4:33-54, 10:20-22, 11:57-61.” Again, these alternative disclosures also do not disclose sufficient structure. RXM-41 (Orchard Decl.) at ¶¶ 145-49. The “step 208” that Complainants identify

(shown at right in Figure 11) is simply a  black box that restates the *function* of “decoding” a current symbol with the current probability model. The ’859 patent admits that step 208 shows a “functional unit.” JXM-2 (’859 Patent) at 13:53-55. The last two passages Complainants cite (10:20-22, 11:57-61) also do not disclose any structure. Both passages merely restate the *function* that during step 208, “the current probability model Pc is updated according to the method described previously in reference to the AC coding.” *Id.* at 10:20-22, 11:57-61.

The first passage Complainants cite (4:33-54) is the same one discussed above that supports Respondents’ construction, i.e., explaining that Arithmetic Coding “updates” a probability model by increasing the occurrence number for the current symbol by 1. *Id.* at 4:42-45. Accordingly, Respondents’ proposed construction for this term should be adopted.

VII. ’556 PATENT

A. “sparse de-noising filter”

Complainants’ Construction	Respondents’ Construction	Staff’s Construction
Plain and ordinary meaning (i.e., a filter that reduces noise based	sparse de-noising filter for performing a second pass to	A filter that reduces noise based on a

on a sparse representation of the signal)	reduce noise, which is not a deringing filter Alternatively, a filter which exploits a sparse image model using an over complete set of linear transforms and hard thresh-holding, which is not a deringing filter	sparse representation of the signal
---	--	-------------------------------------

The parties have two disputes: (1) whether Complainants have disclaimed “deringing filters” from the scope of the ’556 patent, and (2) the construction of a “sparse de-noising filter.”

1. InterDigital Disclaimed “Deringing Filters” During Prosecution

During the prosecution of U.S. Patent No. 9,277,243 (“the ’556 Parent Patent”), from which the ’556 patent continues, InterDigital carved “deringing filters” out from the scope of the claims to distinguish anticipatory prior art. The prosecution history for a limitation in a parent patent is presumed to apply to the identical limitation in subsequent, related patents “unless otherwise compelled.” *Trading Techs. Int’l v. Open E Cry, LLC*, 728 F.3d 1309, 1323 (Fed. Cir. 2013) (citing *Omega Eng’g, Inc. v. Raytek Corp.*, 334 F.3d 1314, 1323–24 (Fed. Cir. 2003)).

The ’556 Parent Patent issued with one claim:

1. A method comprising:

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

RXM-43 (’556 Parent Patent) at claim 1 (emphasis added).

As shown, both the '556 Parent Patent and the asserted '556 patent recite the same limitation requiring an “adaptive sparse de-noising filter for performing a second pass to reduce noise.” *Compare* '556 Parent Patent, at claim 1, *with* JXM-3 ('556 Patent) at claim 1.

Before it was issued, claim 1 of the '556 Parent Patent application required only “two filters” for filtering “a first and a second type of quantization artifact.” This earlier version did not require any specific types of filters or the types of artifacts to be filtered (i.e., did not specify a sparse de-noising filter):

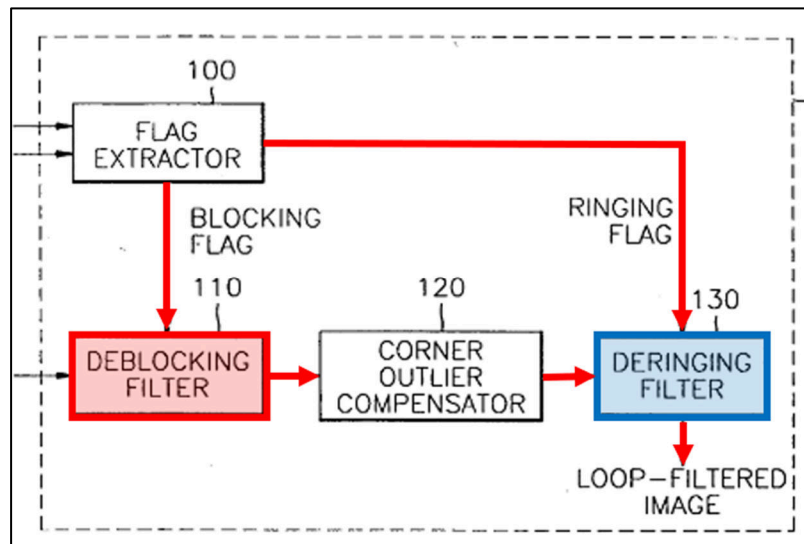
1. (currently amended) An apparatus, comprising: an encoder [[for]] encoding an image region, the encoder having at least two filters for successively performing in-loop filtering to respectively reduce at least a first and a second type of quantization artifact , wherein said encoder renders a decision to selectively enable or disable <u>each of the at least two filters at a given level, the given level being selected from among each of</u> <u>a macroblock level, a slice level, a picture level, and a sequence level.</u>

RXM-22 ('556 Parent Patent File History) at 160019 (highlights added).

The PTO rejected this early claim as anticipated by U.S. Patent Application No. 2003/0219073 (“Lee”) (RXM-42). The examiner found it anticipated because Lee discloses an “encoder having *at least two filters* for successively performing in-loop filtering to respectively reduce at least a first and a second type of quantization artifact.” RXM-22 ('556 Parent Patent File History) at 159847(emphasis added). Specifically, the examiner noted that Lee discloses the two filters as a deblocking filter and deringing filter, citing the following passage in Lee:

According to another aspect of the object, there is provided an apparatus for loop-filtering image data, . . . the apparatus comprising: . . . a *deblocking filter* for deblocking filtering the image data . . . and a *deringing filter* for deringing-filtering . . .

RXM-42 (Lee) at ¶ 31. Lee shows the two deblocking and deringing filters at issue in Figure 1, below:



Id., at Fig. 1 (annotated with the deblocking filter in **red** and deringing filter in **blue**)

In response to the examiner's rejection, the applicant narrowed the claim to specify particular types of filters: "a deblocking filter for performing a first pass to reduce blocking artifacts and a *sparse de-noising filter for performing a second pass* to reduce noise." RXM-22 ('556 Parent Patent File History) at 160112. Specifically, the amended claim recited:

1. (Currently Amended) An apparatus [[,]] comprising:
 an encoder ~~that encodes encoding~~ an image region, the encoder ~~including having~~ at least two filters for successively performing in-loop filtering to ~~respectively~~ reduce at least a first and a second type of quantization artifact, respectively, the at least two filters including a deblocking filter for performing a first pass to reduce blocking artifacts and a sparse de-noising filter for performing a second pass to reduce noise, wherein ~~the~~ [[said]] encoder renders a decision to selectively enable or disable ~~the sparse de-noising filter each of the at least two filters~~ at a given level, the given level being ~~at least selected from among each of~~ a macroblock level, a slice level, a picture level, or [[and]] a sequence level.

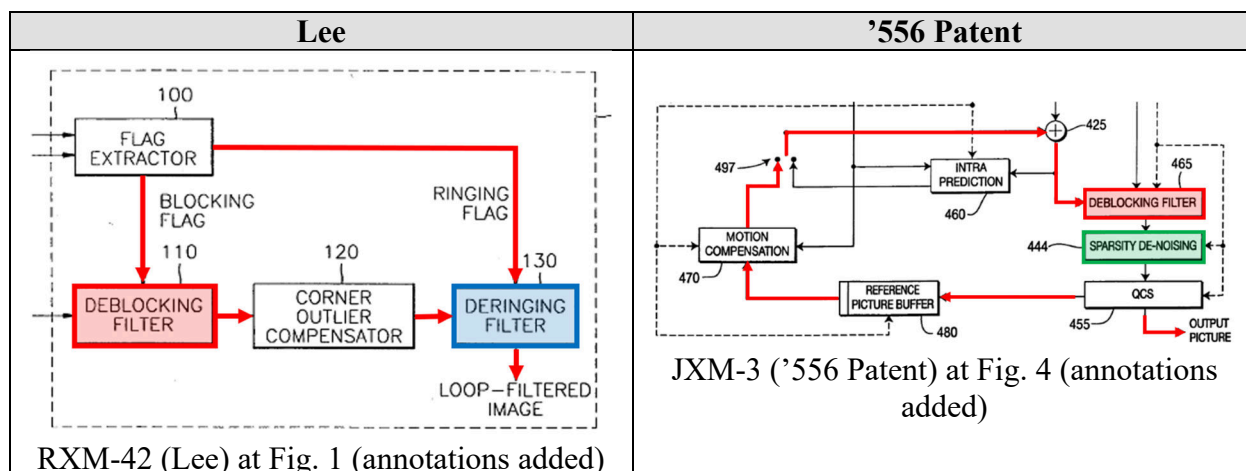
RXM-22 ('556 Parent Patent File History) at 160112 (highlights added, underlining and strikethrough in original). As a result, the claim was narrowed to cover not just any two filters, but specifically a deblocking filter and “sparse de-noising filter.”

Arguing the amended claim language overcame the rejection based on Lee, the applicant clearly and unequivocally distinguished deringing filters (as disclosed in Lee) from sparse de-noising filters (as claimed):

The cited portions of Lee do not expressly or inherently disclose at least a **sparse de-noising filter for performing a second pass to reduce noise**. While the cited portions of Lee may disclose deringing filtering, Lee does not disclose or suggest a **sparse de-noising filter for performing a second pass to reduce noise**.”

RXM-22 ('556 Parent Patent File History) at 160119 (highlighting added, emphasis in original).

To illustrate the applicant’s disclaimer, compare Lee’s filters to the '556 patent’s filters. The key distinction is the applicant’s replacement of Lee’s deringing filter (shown in **blue**) with a sparse de-noising filter (shown in **green**):



The applicant overcame the PTO's rejection by disclaiming deringing filters, arguing specifically that the '556 Parent Patent's sparse de-noising filter is *different from a deringing filter*.

While prosecution disclaimer does not apply if “the alleged disavowal is ambiguous,” that does not apply here because the applicant’s disavowal was “clear and unmistakable” to a POSITA. *Omega Eng’g, Inc. v. Raytek Corp.*, 334 F.3d 1314, 1326 (Fed. Cir. 2003). For example, prosecution disclaimer does not apply if the applicant simply describes features of the prior art and does not distinguish the claimed invention based on those features. *Computer Docking Station Corp. v. Dell, Inc.*, 519 F.3d 1366, 1374–75 (Fed. Cir. 2008); *Grober v. Mako Prods.*, 686 F.3d 1335, 1342 (Fed. Cir. 2012). But here, the applicant did not merely describe features of the prior art; the applicant very intentionally *distinguished* the '556 Parent Patent from Lee by explaining that the '556 Parent Patent claims a sparse de-noising filter, *not a deringing filter*. See RXM-22 ('556 Parent Patent File History) at 160119 (arguing that “[w]hile the cited portions of Lee may disclose deringing filtering, Lee does not disclose or suggest a sparse de-noising filter”).

The applicant repeated its disclaimer of deringing filters throughout the lineage of the patent family, for example in U.S. Patent No. 11,089,337 (“the '337 patent”), which is a continuation of the '556 patent. In response to another obviousness rejection over Lee, the

applicant again distinguished Lee’s deringing filters, stating that “Lee discloses a ringing denoising filter and *does not disclose a sparse denoising filter.*” RXM-21 (’337 File History) at 159369 (quoting the examiner’s July 2, 2018 Office Action, at 8 (emphasis added)); *id.* at 159315. Both the applicant and the examiner made this distinction repeatedly throughout the prosecution history. *See, e.g.*, RXM-21 at 159380 (Office Action) (quoting same), 159409 (Appeal Brief) (quoting same). The prosecution history reveals a clear intent to narrow the scope of the claimed invention to disclaim deringing filters, thereby carving out swaths of prior art, like Lee, which disclose deringing filters.

The applicant’s disclaimer of deringing filters was not limited to the ’556 patent’s U.S. family but also extends to its foreign counterparts. As background, the ’556 patent continues from international application PCT/US2007022795 (“the PCT ’795 application”). During the European prosecution of the PCT ’795 application, which was assigned European Patent Application No. 07970824.5 (“the EP ’824 application”), the applicant made precisely the same amendment to overcome principally the same rejection.

During prosecution, claim 1 of the EP ’824 application read as follows:

Claims

5 1. An apparatus comprising an encoder (300) for encoding an
image region, the encoder **having at least two filters**
(365, 344) for successively performing in-loop filtering
to respectively reduce at least **a first and a second type**
of quantization artifact, wherein each filter after the
10 first filter is adapted to correct a particular type of
artifact.

RXM-23 (EP ’824 App. File History) at 160527. Note the absence of any specific limitations regarding the types of filters or the types of artifacts to be filtered (i.e., no limitation specifying a

deblocking or sparse de-noising filter), equivalent to the independent claims of the '556 Parent Patent as originally drafted.

The EPO objected to the claims because “it is not defined which filters are comprised in the two-pass filter and, in particular, it is not defined in technical terms which pass is performed by each filter and how it is performed.” RXM-23 (EP '824 App. File History) at 160541. As a result, and just as done in the U.S. '556 Parent Patent application, the applicant amended the independent claim to “further specify the type of filters.” *Id.* at 160584. The amended claim was narrowed as follows:

Claims (with revision marks) – Request #2

1. An apparatus, comprising an encoder (300) ~~for encoding an image region,~~
wherein the encoder comprises:
- a deblocking filter for performing a first pass in-loop filtering on a
reconstructed image region to obtain a filtered image region; and
- a sparse filter for performing a second pass in-loop filtering on said filtered
image region.
~~_having at least two filters (365, 344) for successively performing in-loop~~
~~filtering to respectively reduce at least a first and a second type of quantization~~
~~artifact, wherein each filter after the first filter is adapted to correct a particular type of~~
~~artifact.~~

Id. at 160593.

The applicant argued that this amended language overcame the EPO’s rejection over an article by Z. Fanz¹³ (referred to in the prosecution history and herein as “D3”) because “D3 solves a different problem, namely *blocking and ringing artifacts*.” RXM-23 (EP '824 App. File History)

¹³ Fanz Z., et al., “JPEG decompression with reduced artifacts,” Journal of the American Veterinary Medical Association, vol. 2186, 9 February 1994 (1994-02-09), pages 50-55.

at 160664 (emphasis added). The applicant noted that “none of the cited references teaches having two in-loop filters the first one being a deblocking filter *and the second one being a sparse filter.*” *Id.* at 160584 (emphasis added). Like its U.S. counterparts, the applicant distinguished the invention’s use of sparse denoising filters from deringing filters, which again were pervasive in the prior art.

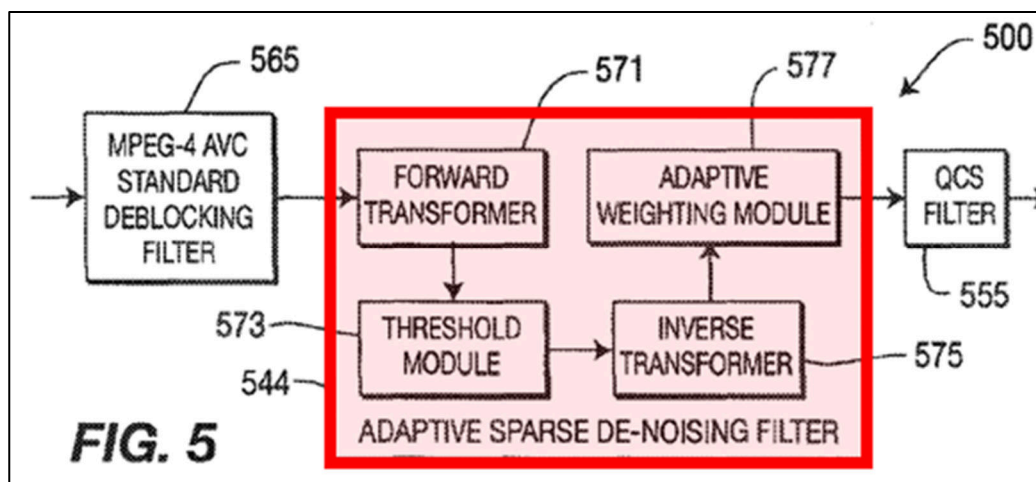
Prosecution disclaimer “promotes the public notice function of the intrinsic evidence and protects the public’s reliance on definitive statements made during prosecution.” *Omega Eng’g, Inc. v. Raytek Corp.*, 334 F.3d 1314, 1324 (Fed. Cir. 2003). Complainants’ position in this litigation undermines this public notice function, as they attempt to recapture the claim scope the applicants disclaimed during prosecution. Specifically, Complainants allege the “sparse denoising filter” of the ’556 patent is met by the “CDEF Process” of the AV1 Decoding Standard. *See* RXM-44 (First Supplemental Responses of Complaints to Respondents’ First Set of Interrogatories, Ex. 5) at 10–24. RXM-44. However, as Complainants’ own infringement contentions show, AV1’s CDEF process is a *deringing* filter. *Id.* at 11–16 (“The purpose of CDEF is to perform *deringing*”; “cdef_damping_minus_3 controls the amount of damping *in the deringing filter*”) (emphasis added). The AV1 standard does not use the term “sparse” or “sparsity” in its description of the CDEF process, nor anywhere in its 681-page disclosure. *See, generally*, Complaint, Ex. 40. That is because the CDEF process is a deringing filter, not a sparse de-noising filter. Complainants should be held to the disclaimer the applicant repeatedly made during prosecution.

2. Alternatively, Sparse De-noising Filter Should Be Construed According to the Specification

To the extent the ALJ deems it necessary to construe “sparse de-noising filter” beyond simply reflecting the disclaimer, the construction should reflect the definition provided by the

inventors: a filter “which exploits a sparse image model using an over complete set of linear transforms and hard thresh-holding.” JXM-3 (’556 Patent) at 2:8—10. Respondents’ proposed construction adopts the specification’s definition and, as described below, is further consistent with the applicant’s own representations during prosecution of related applications. Complainants’ proposed construction for “sparse de-noising filter” is not in the specification of the ’556 patent and contradicts the applicant’s own statements made in the prosecution of the patent family.

The ’556 patent repeatedly and consistently describes the sparse de-noising filter as a filter that exploits a sparse image model using an over complete set of linear transforms and hard thresh-holding. In addition to the express description at 2:08–10, the inventors explain the sparse de-noising filter creates numerous de-noised samples using “three steps: transform; transform coefficients threshold; and inverse transform.” JXM-3 (’556 Patent) at 2:14–16. Then, then de-noised samples are “combined using weighted averaging at every pixel.” *Id.* 2:16–21. The process is illustrated in Figure 5:



Id. at Fig. 5.

Later, the inventors detail each individual step. First, the sparse de-noising filter determines “an over *complete transform* set H_i , $i=1, \dots, M$.” JXM-3 (’556 Patent) at 13:24–25

(emphasis added). In the next step, the filter determines “a *threshold* γ .” *Id.* at 13:36. Finally, the sparse de-noising filter applies “the adaptive weighting method to combine the outcome of the inversed transforms.” *Id.* at 13:43–44.

At each disclosure of a sparse de-noising filter the inventors repeatedly describe it as performing the same series of steps, “includ[ing] a forward *transformer* 571, . . . a *threshold* module 573, . . . an *inverse transformer* 575, . . . and an adaptive weighting module 577.” *Id.* at 11:17–24. The specification of the ’556 patent does not describe or contemplate any embodiment of an “adaptive sparse de-noising filter” not conforming to this model.

Complainants’ construction is also inconsistent with the representations made by InterDigital during the prosecution of EP ’824 application. There, the EPO noted two possible interpretations of “sparse filter.” The first “seeming straightforward interpretation” was “a filter represented as a matrix of coefficients containing very view non-zero elements.” The EPO noted that “[t]his interpretation appears to be the one *highlighted by the appellant* at the oral proceedings before the examining division.” RXM-23 (EP ’824 App. File History) at 160890 (emphasis added). The EPO then noted that the “[t]his interpretation is however inconsistent with” how the specification itself defined a “sparse filter.” *Id.* The EPO noted that according to the specification, “the term ‘sparse’ is not intended to relate to a matrix of filtering coefficients but to a representation of an image region using an over complete set of linear transforms.” *Id.*

In response, InterDigital amended the claims of the EP ’824 application to reflect the definition set forth in the specification:

Claims (Draft Version) – Request #3

1. An apparatus, comprising an encoder (300), wherein the encoder comprises:
- a deblocking filter for performing a first pass in-loop filtering on a reconstructed
 - 5 image region to obtain a filtered image region; and
 - an adaptive sparse de-noising filter which exploits a sparse image model using an over complete set of linear transforms for performing a second pass in-loop filtering on said filtered image region.

RXM-23 (EP '824 App. File History) at 160980.

To show the amendment did not introduce new subject matter, InterDigital identified disclosures at page 19 line 30 (“In an embodiment, an adaptive sparse de-noising filter operation involves the following steps. In one step, an over complete transform set H_i , $i = 1, \dots, M$ is determined”) and page 2 lines 25-26 (“a nonlinear de-noising filter [...] which exploits a sparse image model using an overcomplete set of linear transforms”), language identical to that in the specification of the asserted '556 patent. *Id.* at 160973. Ultimately, InterDigital submitted the amended claim “further specifying that the adaptive sparse denoising filter is based on *a sparse image model using an over complete set of linear transforms . . .*” *Id.* (emphasis in original).

Contrary to InterDigital’s representations to the European Patent Office, Complainants now propose construing the term to mean a “filter that reduces noise based on a sparse representation of the signal.” The '556 patent does not provide any support for this construction. The patent does not define or discuss a “sparse representation of the signal.” It would not be clear to a person of skill in the art whether the term “sparse representation of the signal” would be confined to algorithms involving transforms and hard-thresholding, or if it more generally refers to any representation of a signal that conveys less than the complete set of elements or data. RXM-41 (Orchard Decl.) at ¶ 153. Additionally, the claim term “reduces noise” does not clearly

establish the metes and bounds of the scope of the claims: it is silent as the extent of noise reduction that would be necessary to satisfy the claim element, and also fails to define the specific types of noise to be filtered (e.g., quantization noise, gaussian noise, or specific artifacts such as blocky or ringing artifacts). RXM-41 (Orchard Decl.) at ¶ 154. Lastly, the specification does not provide any detail on what kinds of functionality would constitute filtering “based on” the sparse representation of the signal. For example, it is not clear whether the specific values in the sparse representation must be altered or whether simply extracting information from a sparse representation (or making certain inferences based on the sparse representation) to inform some action taken on the image itself would also constitute infringement. Complainants’ proposal to introduce the language “based on” creates ambiguity within the claim term, and thus does not provide the public a clear understanding of the metes and bounds of the claim.

B. “adaptive sparse de-noising filter”

Complainants’ Construction	Respondents’ Construction	Staff’s Construction
Plain and ordinary meaning (i.e., a sparse denoising filter (a filter that reduces noise based on a sparse representation of the signal) that can adjust its own parameters according to the image region)	sparse de-noising filter for performing a second pass to reduce noise, which is not a deringing filter, that can adjust its own parameters automatically; Alternatively, a filter which exploits a sparse image model using an over complete set of linear transforms and hard thresh-holding, which is not a deringing filter, that can adjust its own parameters automatically.	A sparse denoising filter (a filter that reduces noise based on a sparse representation of the signal) that can adjust its own parameters automatically

The parties dispute “adaptive” should be construed to mean “that can adjust its own parameters automatically,” or more narrowly as “that can adjust its own parameters according to the image region.”

Respondents and Staff agree the proper construction is the former, as it is supported by the specification of the '556 patent: “The nonlinear de-noising filter *automatically* becomes high-pass, or low-pass, or band-pass, and so forth.” JXM-3 ('556 Patent) at 2:10–14 (emphasis added).

Complainants' proposed construction unduly narrow and does not encapsulate the broad range of adaptability described in the '556 patent or known by POSITAs. As used in the '556 patent, the term “adaptive” encompasses adjustments to parameters based on a variety of inputs, not limited to the image region alone. The patent explicitly describes how a sparse de-noising filter can adapt based on, for example, “non-stationary image statistics,” *id.* at 2:04–09, and the “number of non-zero coefficients” in the sparse representation, *id.* at 13:53–56. Furthermore, the patent details how the sparse de-noising filter's threshold level (γ) can be automatically set based on inputs like the “quantization parameter as well as to the sequence content.” *Id.* at 11:64–66. The patent describes that the threshold level can be set at different levels of abstraction, such as slice, picture or frame level. Conceivably, if the threshold level is set at the picture or frame level, the filter would not be adjusting to any aspect of the image “region” specifically, but might instead be uniformly applied across the image. *Id.* at 12:01–35. Moreover, a POSITA could readily conceive of various other inputs that might influence an adaptive sparse de-noising filter to automatically adjust its own parameters, which the patent explicitly acknowledges. *Id.* at 12:39–42.

Constraining the term “adaptive” to adjustments based solely on image region not only misinterprets the patent's broader technical context but also unduly narrows its scope. This narrow interpretation fails to account for the patent's comprehensive description of the various inputs on which it bases its parameters. Staff and Respondents' construction, on the other hand, emphasizes that the sparse de-noising filter is adaptive in nature because it automatically adjusts various

parameters without the need for specific intervention, while not arbitrarily limiting the inputs on which the filter basis its adaptive behavior. Such a construction is more accurately aligned with the technological realities of the sparse de-noising filter, which may be designed to respond to a wide array of inputs, as described in the patent and understood by those skilled in the art. *See* RXM-41 (Orchard Decl.) at ¶¶ 156–57.

VIII. CONCLUSION

For the reasons discussed in greater detail below, Respondents respectfully request that the ALJ adopt their proposed construction of each term.

Dated: January 26, 2024

Respectfully submitted,

FISH & RICHARDSON P.C.

By: /s/ Linhong Zhang

Adam R. Shartzter
Ruffin B. Cordell
Michael J. McKeon
Richard A. Sterba
Christian Chu
Linhong Zhang
Sunyoung Park
John Thuermer
Joshua Carrigan
Daniel Lee

Fish & Richardson P.C.

1000 Maine Avenue, S.W., Suite 1000
Washington, D.C. 20024
Telephone: (202) 783-5070

David Barkan

Leeron G. Kalay

Fish & Richardson P.C.

500 Arguello Street, Suite 400
Redwood, CA 94063
Telephone: (650) 839-5070

Scott M. Flanz

Kyle Fleming

Allison E. Elkman
Fish & Richardson P.C.
7 Times Square
20th Floor
New York, NY 10036
Telephone: (212)765-5070

Matthew Johnson
Fish & Richardson P.C.
60 South 6th Street, Suite 3200
Minneapolis, MN 55402
Telephone: (612) 335-5070

Counsel for Respondents
Lenovo PC HK Limited, Lenovo (United
States) Inc., and Motorola Mobility LLC

CERTIFICATE OF SERVICE

I hereby certify that true and correct copies of the foregoing document:

RESPONDENTS' OPENING MARKMAN BRIEF

have been filed and served on this 26th day of January, 2024, on the following:

The Honorable Lisa R. Barton Secretary to the Commission U.S. International Trade Commission 500 E Street, S.W. Washington, D.C. 20436	☐ Via First Class Mail ☐ Via Hand Delivery ☐ Via Federal Express ☒ Via Electronic Filing
The Honorable MaryJoan McNamara Administrative Law Judge U.S. International Trade Commission 500 E Street, S.W., Room 317 Washington, D.C. 20436 Email: mcnamara337@usitc.gov	☐ Via First Class Mail ☐ Via Hand Delivery (two double sided paper copies) ☐ Via Federal Express ☒ Via Electronic Mail
Jennifer Dienes, Esq. Investigative Attorney Office of Unfair Import Investigations U.S. International Trade Commission 500 E Street, S.W. Washington, D.C. 20436 Email: jennifer.dienes@usitc.gov	☐ Via First Class Mail ☐ Via Hand Delivery ☐ Via Federal Express ☒ Via Electronic Mail ☒ Via USITC Box Platform
Adam D. Swain, Esq. Alston & Bird LLP 950 F Street NW Washington, D.C. 20004 Email: IDG.Lenovo.ITC@alston.com <i>Counsel for Complainants InterDigital Inc., InterDigital VC Holdings, Inc., InterDigital Patent Holdings, Inc., and InterDigital Madison Patent Holdings SAS</i>	☐ Via First Class Mail ☐ Via Hand Delivery ☐ Via Federal Express ☒ Via Electronic Mail

/s/ Ashley Cox