**UNITED STATES PATENT AND TRADEMARK OFFICE**

**BEFORE THE PATENT TRIAL AND APPEAL BOARD**

---

Samsung Electronics America, Inc.,
Petitioner

v.

Prisua Engineering Corp.,
Patent Owner.

---

Patent No. 8,650,591 to Prieto

IPR Case No. IPR2017-01188

---

# PETITION FOR *INTER PARTES* REVIEW OF CLAIMS 1-4, 8 AND 11 OF U.S. PATENT 8,650,591 UNDER 35 U.S.C. §§ 311-319 AND 37 C.F.R. §§ 42.100 ET SEQ.

---

## TABLE OF CONTENTS

TABLE OF EXHIBITS ..... iv

I. MANDATORY NOTICES UNDER 37 C.F.R. § 42.8(a)(1) FOR *INTER PARTES* REVIEW ..... 1

A. Real Party in Interest Under 37 C.F.R. § 42.8(b)(1) ..... 1

B. Related Matters Under 37 C.F.R. § 42.8(b)(2) ..... 1

C. Lead and Backup Counsel Under 37 C.F.R. § 42.8(b)(3) and Service Information under 37 C.F.R. § 42.8(b)(4) ..... 2

II. PAYMENT OF FEES UNDER 37 C.F.R. § 42.15 ..... 2

III. Certification of Word Count Under 37 C.F.R. § 42.24(d) ..... 3

IV. REQUIREMENTS FOR IPR UNDER 37 C.F.R. § 42.104 ..... 3

A. Grounds for Standing Under 37 C.F.R. § 42.104(a) ..... 3

B. Identification of Challenge Under 3 C.F.R. § 42.104(b) and Relief Requested ..... 3

V. SUMMARY OF THE '591 PATENT ..... 5

A. Overview of the Alleged Invention ..... 5

B. The Prosecution History ..... 6

C. The Challenged Claims ..... 8

VI. RELEVANT INFORMATION CONCERNING THE CONTESTED PATENT ..... 8

A. Person of Ordinary Skill in the Art ..... 8

B. Claim Construction ..... 8

1. "digital extraction" ..... 9

VII. SPECIFIC GROUNDS FOR PETITION ..... 11

A. Ground 1: Claims 1, 2, 8, and 11 are anticipated by or rendered obvious by Senftner ..... 11

1. Overview of Senftner ..... 11

2. Claims ..... 13

B. Ground 2: Claims 3 and 4 are obvious over Senftner in view of Levoy ..... 38
1. Scope and Content of the Prior Art ..... 38
2. Motivation to Combine ..... 41
3. Claims ..... 43
C. Ground 3: Claims 1, 2, 8, and 11 are obvious under Sitrick ..... 46
1. Overview of Sitrick ..... 46
2. Claims ..... 48
D. Ground 4: Claims 3 and 4 are rendered obvious by Sitrick in view of Levoy ..... 66
1. Scope and Content of the Prior Art ..... 66
2. Motivation to Combine ..... 67
3. Claims ..... 68
VIII. CONCLUSION ..... 70
CERTIFICATE OF SERVICE ..... 72

## TABLE OF EXHIBITS

| Exhibit | Description |
|---|---|
| 1001 | U.S. Patent No. 8,650,591 ("'591 patent") |
| 1002 | File history of U.S. Patent No. 8,650,591 |
| 1003 | Declaration of Edward Delp Regarding U.S. Patent No. 8,650,591 |
| 1004 | *Prisua Engineering Corp. v. Samsung Electronics Co., Ltd., et al.*, CA No. 1:16-cv-21761-KMM, Deposition Transcript of Dr. Yolanda Prieto (January 17, 2017) |
| 1005 | *Prisua Engineering Corp. v. Samsung Electronics Co., Ltd., et al.,* CA No. 1:16-cv-21761-KMM, Joint Claim Construction and Prehearing Statement (November 21, 2016) |
| 1006 | U.S. Patent No. 7,460,731 to Senftner et al. ("Senftner") |
| 1007 | U.S. Patent Application Publication No. 2005/0151743 to Sitrick ("Sitrick") |
| 1008 | U.S. Patent Application Publication No. 2009/0309990 to Levoy et al. ("Levoy") |
| 1009 | Negahdaripour Decl. ISO Patent Owner's Opening Claim Construction Brief ("Negahdaripour Decl.") |
| 1010 | U.S. Patent Application Publication No. 2006/0097991 |
| 1011 | U.S. Patent No. 7,307,623 to Enomoto |
| 1012 | U.S. Patent No. 4,686,332 to Greanias et al. |
| 1013 | Edward Delp Decl. ISO Petitioner's Responsive Claim Construction Brief |
| 1014 | U.S. Patent Application Publication No. 2008/0148167 to Zeev Russak et al. |

| | |
|---|---|
| 1015 | MPEG \| The Moving Picture Experts Group website, (see http://mpeg.chiariglione.org/) |

Samsung Electronics America, Inc. ("Petitioner" or "Samsung") hereby seeks *inter partes* review of Claims 1-4, 8, and 11 ("the Challenged Claims") of U.S. Patent No. 8,650,591. (Ex. 1001 (the "'591 patent").) The Challenged Claims of the '591 patent do not claim anything inventive. The Challenged Claims disclose substituting an extracted image from one video data stream into another to create an edited video. However, prior to the '591 patent, substitution of images extracted from one video into another video was well-known, and the Challenged Claims should be canceled for the reasons described in this Petition.

## I. MANDATORY NOTICES UNDER 37 C.F.R. § 42.8(a)(1) FOR *INTER PARTES* REVIEW

### A. Real Party in Interest Under 37 C.F.R. § 42.8(b)(1)

The real parties in interest are Samsung Electronics Co., Ltd., 416 Maetan-3dong, Yeongton-gu, Suwon-City, Gyeonggi-do 443-742, South Korea; Samsung Electronics America, Inc., 85 Challenger Road, Ridgefield Park, New Jersey 07660; and Samsung Electronics Latinoamerica Miami, Inc., 9850 NW 41st St #350, Doral, FL 33178.

### B. Related Matters Under 37 C.F.R. § 42.8(b)(2)

A pending federal district court litigation may be affected by the decision in this proceeding: *Prisua Engineering Corp. v. Samsung Electronics Co., Ltd et al.*, No. 1:16-cv-21761-MOORE (S.D. Fla. May 17, 2016) ("the Litigation"). The

undersigned is unaware of any other matter that would affect, or be affected by, a decision in the proceeding.

### C. Lead and Backup Counsel Under 37 C.F.R. § 42.8(b)(3) and Service Information under 37 C.F.R. § 42.8(b)(4)

Petitioner designates the following lead and backup counsel:

| | |
|---|---|
| Lead Counsel: | Heath J. Briggs (Reg. No. 54,919)<br>Greenberg Traurig, LLP<br>1200 17th St., Suite 2400<br>Denver, CO 80202<br>Telephone: 303-685-7418<br>Facsimile: 720-904-6118<br>BriggsH@gtlaw.com |
| Backup Counsel: | Patrick J. McCarthy (Reg. No. 62,762)<br>Greenberg Traurig, LLP<br>2101 L street, N.W., Suite 1000<br>Washington, D.C. 20037<br>Telephone: 202-533-2386<br>Facsimile:202-331-3101<br>McCarthyP@gtlaw.com |

Service on Petitioner may be made by mail or hand delivery to: Greenberg Traurig, LLP, 1200 17th St., Suite 2400, Denver, CO, 80202. Petitioner also consents to electronic service by emailing counsel of record at briggsh@gtlaw.com, mccarthyp@gtlaw.com and PrisuaGTIPR@gtlaw.com.

## II. PAYMENT OF FEES UNDER 37 C.F.R. § 42.15

Petitioner authorizes the U.S. Patent & Trademark Office to charge Deposit Account No. 50-2638 for the fee set in 37 C.F.R. § 42.15(a) for this Petition and further authorizes for any additional fees to be charged to this Deposit Account.

## III. CERTIFICATION OF WORD COUNT UNDER 37 C.F.R. § 42.24(d)

Petitioner certifies that the word count in this Petition is XX words, as counted by the word-processing program (Microsoft Word 2010) used to generate this Petition, where such word count excludes the table of contents, table of authorities, mandatory notices, certificate of service, appendix of exhibits, and this certificate of word count. This Petition is in compliance with the 14,000 word limit set forth in 37 C.F.R. § 42.24(a)(1)(i).

## IV. REQUIREMENTS FOR IPR UNDER 37 C.F.R. § 42.104

### A. Grounds for Standing Under 37 C.F.R. § 42.104(a)

Petitioner certifies that the '591 patent is available for *inter partes* review, and that Petitioner is not barred or estopped from requesting an *inter partes* review on the grounds identified in the petition.

### B. Identification of Challenge Under 3 C.F.R. § 42.104(b) and Relief Requested

Petitioner respectfully requests that claims 1-4, 8, and 11 of the '591 patent be cancelled based on the following grounds of unpatentability, explained in detail below.

| Ground of Unpatentability | '591 Patent Claim(s) | Basis for Rejection |
|---|---|---|
| **Ground 1** | 1, 2, 8 and 11 | Anticipated by or rendered obvious by U.S. Patent No. 7,460,731 ("Senftner") |
| **Ground 2** | 3 and 4 | Rendered obvious by Senftner in view of U.S. Patent Application Publication No. 2009/0309990 to Levoy et al. ("Levoy") |
| **Ground 3** | 1, 2, 8, and 11 | Rendered obvious by U.S. Patent Application Publication No. 2005/0151743 ("Sitrick") |
| **Ground 4** | 3 and 4 | Rendered obvious by Sitrick in view of Levoy |

The '591 patent issued on February 11, 2014, from Application No. 13/042,955 filed on March 8, 2011. The '591 patent claims priority to provisional application no. 61/311,892 filed on March 9, 2010.[1]

U.S. Patent No. 7,460,731 ("Senftner") qualifies as prior art under at least pre-AIA 35 U.S.C. §102(b). Senftner issued on December 2, 2008, more than one year before the earliest possible priority date of the '591 patent.

U.S. Patent Application Publication No. 2005/0151743 ("Sitrick") qualifies as prior art under at least pre-AIA 35 U.S.C. §102(b). Sitrick was published on July 14, 2005, more than one year prior to the earliest possible priority date of the '591 patent.

[1] Petitioner does not concede the '591 patent is entitled to the priority date of this provisional patent application.

U.S. Patent Application Publication No. 2009/0309990 ("Levoy") qualifies as prior art under at least pre-AIA 35 U.S.C. §§102(a),(e). Levoy was filed on June 11, 2008 and published on December 17, 2009, both of which occurred prior to the earliest possible priority date of the '591 patent.

## V. SUMMARY OF THE '591 PATENT

### A. Overview of the Alleged Invention

The '591 patent is entitled "Video Enabled Digital Devices for Embedding User Data in Interactive Applications" and issued from an application filed on March 8, 2011.

The '591 patent is directed to a system that creates a new composite video by substituting a portion of an original video data stream with an image from a user input video data stream. FIG. 3 (reproduced below) shows the preferred "image substitution" described by the patent. According to the specification, "a user input 150 of a photo image of the user used to *replace the face of the image* shown on the device 108. The user transmits the photo image 150 by wired or wireless means to the device 108. The *image substitution* is performed and the device 108 shows the substituted image 190." (Ex. 1001 at 2:66-3:4 (emphasis added).)

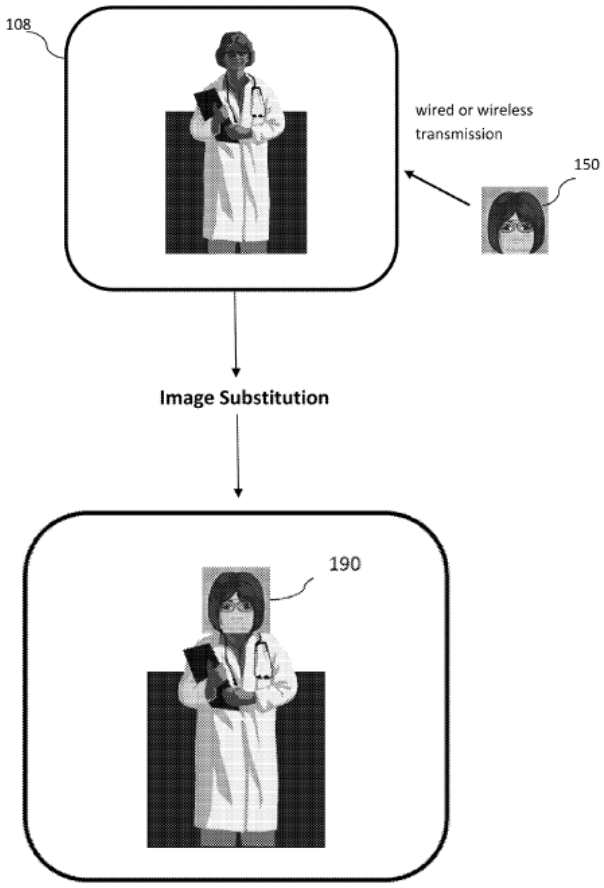



***FIG. 3***

(*Id.* at Fig. 3.)

**B. The Prosecution History**

The original claims of the '591 patent were rejected by the patent examiner who found the claims unpatentable under § 101, as well as anticipated by U.S. Patent No. 8,024,768 to Berger ("Berger"). (*Id*. at 46.)

In a response filed March 13, 2013, applicant cancelled all original claims, and added new claims 21-40 which were essentially redrafted versions of the original claims. (*Id*. at 58.)

On June 5, 2013, the examiner issued a final rejection, finding all claims anticipated based on Berger. (*Id*. at 77.)

On July 16, 2013, applicant faxed a telephonic interview request to the examiner stating that applicant would like to discuss "Berger's failure to teach video/image content substitution as required by the independent claims." *(Id*. at 89.) A telephone interview was held on July 22, 2013. (*Id*. at 97.)[2]

In a response filed August 2, 2013, applicant cancelled claims 21-26, added new claims 41-47 (*Id*. at 97.)

In an Advisory Action mailed August 21, 2013, the examiner rejected all claims noting that, despite applicants claim amendments, "Berger still meets the claim limitations". (*Id*. at 106.)

In a response filed September 5, 2013, applicant cancelled all prior pending claims, added new claims and argued the new claims were patentable over Berger because, *inter alia*, "Berger does not perform a substitution of one selected image for another selected image. Berger does not discuss a substitution at all. Instead,

[2] A USPTO Interview Summary form was not found in the record in PAIR. The only mention of the July 22, 2013 telephonic interview is in applicant's response of August 2, 2013.

Berger generates a new video as a result of resizing, cropping, and pan-and-scan." (*Id*. at 125.) A notice of allowance was subsequently mailed. The reasons for allowance were, in essence, that the prior art failed to disclose the features of the allowed claims. (*Id*. at 134-137.)

The prior art cited in this Petition were not analyzed, considered, described, or cited during prosecution of the '591 patent.

### C. The Challenged Claims

Claims 1 and 11 are independent claims in the '591 patent. Claims 2-4 and 8 depend directly or indirectly from Claim 1.

## VI. RELEVANT INFORMATION CONCERNING THE CONTESTED PATENT

### A. Person of Ordinary Skill in the Art

A POSITA in the field of the '591 patent at the time of the earliest possible priority date (March 9, 2010) would be at least an engineer with a Bachelor of Science degree and at least three years of imaging and signal processing experience or would have earned a Master's Degree in Electrical Engineering and at least two years of professional experience in signal, image, and video processing. (Ex. 1003 at ¶25.)

### B. Claim Construction

Pursuant to 37 C.F.R. § 42.100(b), the claims in *inter partes* review are given the "broadest reasonable construction in light of the specification". For the

purposes of this proceeding, the following claim term should be construed as set forth below. For the purposes of this proceeding, Petitioner requests that each of the remaining claim terms be given their plain meaning.[3]

**1. "digital extraction"**

Claims 1-4 and 11 of the '591 patent use the terms "digitally extracted," "digital extraction," and like terms in the context of creating the first or second images from pixel(s) of a video data stream. (Ex. 1001 at Cl. 1 (7:14-20, 39-40, 45) and Cl. 8 (8:28-34, 49-50, and 55-56).) The specification of the '591 patent never uses the word extraction. Instead, the '591 patent uses the word "substitution":

> "*Summary of the Invention*: Briefly, according to an embodiment of the invention a method for generating an edited video data stream from an original video stream wherein generation of said edited video stream comprises a step of: *substituting at least one object in a plurality of objects in said original video stream by at least a different object*. According to another embodiment, an interactive broadcast TV system enhances the user experience by allowing said user the

[3] In the Litigation, Petitioner proposed construction of various terms in accordance with the *Phillips* standard. (Ex. 1005 at 5-22.) Petitioner maintains that those proposed constructions are correct in the Litigation under the *Phillips* standard.

> capability of *providing his/her object data for use in the substitution of an object data in a broadcast video by said user object data*." (Ex. 1001 at 1:40-51 (emphasis added).)

FIG. 3 illustrates this "substitution" methodology, and specifically states "Image Substitution". In fact, the '591 uses the word substitution twenty-seven (27) times in its specification and without explaining how pixel(s) are "digitally extracted" from a data stream to form an image.

Further, during prosecution, the term "digitally extracted" was not explained. (*See* Ex. 1002 at 125-126.)

In the Litigation, Patent Owner proposed "extracting" to mean "select and separate out". (Ex. 1005 at 16-17.) Patent Owner's expert, in arguing that "extraction" does not require a complete removal of that selected data from the original data, testified:

> "In sum, one of ordinary skill in the art attempting to practice the '591 patent's disclosure would likely simply "select" and "copy out" into memory the information present in the data streams, and generate a new third "data stream" from memory. Prisua's proposed construction therefore comports with the understanding of the term as known to those of skill in the art." (Ex. 1009 at ¶53; *see also, id.* at 48-52.)

Based on the foregoing and in view of Patent Owner's assertion of its claims in the Litigation, Petitioner proposes that the broadest reasonable interpretation of "digitally extracted," "digital extraction," and like is "to digitally select and separate out, such as by copying."

## VII. SPECIFIC GROUNDS FOR PETITION

### A. Ground 1: Claims 1, 2, 8, and 11 are anticipated by or rendered obvious by Senftner

#### 1. Overview of Senftner

Senftner is entitled "Personalizing a video" and is directed to creating a personalized video through partial image replacement in the original video. (Ex. 1006 at Abstract.) Like the '591 patent, Senftner discloses computerized systems and methods for replacing images or video of an "original actor" or "target" in an original video with images or video of a "new actor" or "target replacement." (*See, e.g., id*. at 2:33-54, 9:25-31.) In particular, Senftner explains:

> "The creation of personalized video is a combination of multiple fields that in totality *allow for the alteration of video sequences such that individuals are able to replace the participants of an original video with themselves, their friends, their family members or any individuals, real or imagined, which they have images depicting*. (*Id*. at 5:42-47 (emphasis added).)

FIG. 1 of Senftner, reproduced below, provides a high-level overview of Senftner's image substitution process.

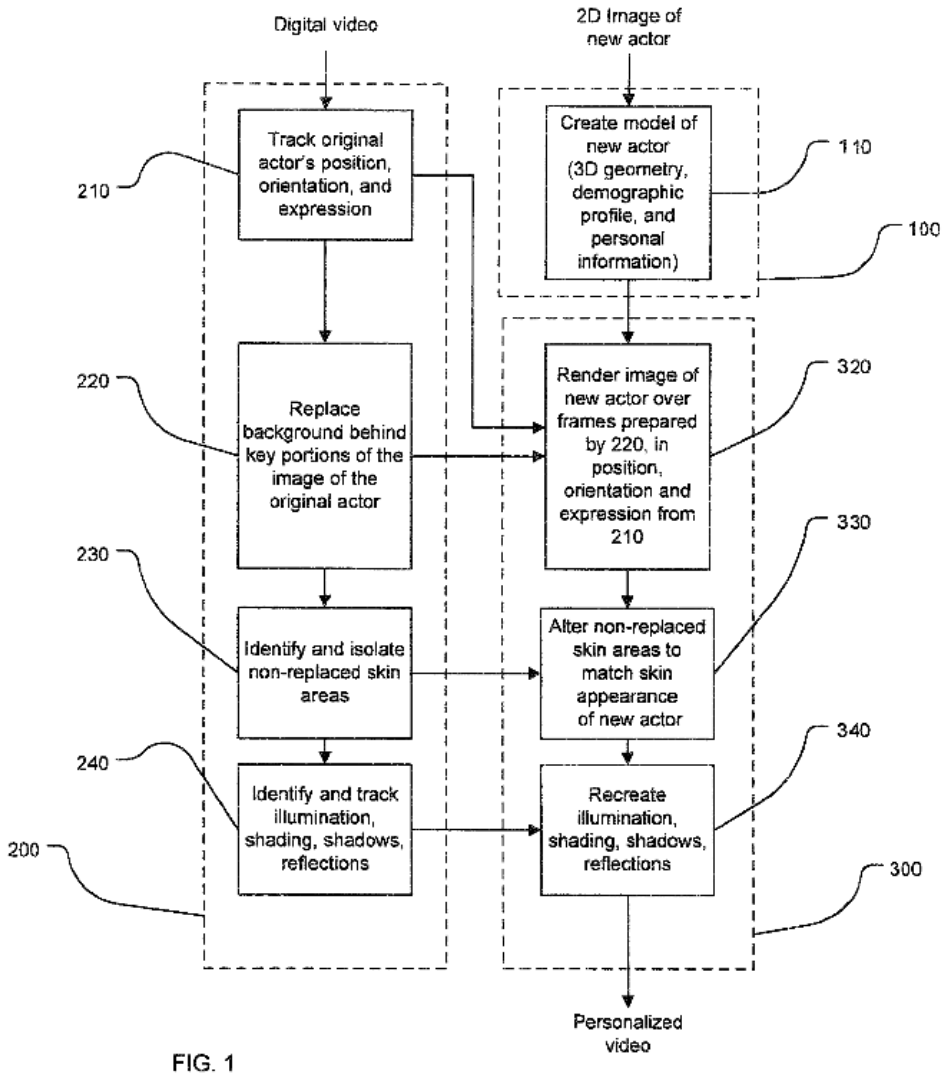



FIG. 1

As Senftner explains, process 100 of FIG. 1 relates to obtaining images of the new actor (called the "actor modeling process"), process 200 relates to preparing an original video having the original actor for substitution (called the "video preparation process"), and process 300 relates to making the personalized video based on steps 100 and 200 (called the "personalization process"). (*Id*. at 9:32-52.)

In order to facilitate the substitution, motion correction may be applied to one of or both of the original actor and the new actor during capture of the images or video. (*Id*. at 6:10-14, 17:10-23.)

For the reasons provided below, Senftner anticipates or renders obvious claims 1, 2, 8 and 11 of the Challenged Claims.

**2. Claims**

**[1-PREAMBLE-i] An interactive media apparatus for generating a displayable edited video data stream from an original video data stream,**

Senftner discloses this limitation. Senftner discloses "personalizing video through partial image replacement". (*Id.* at Abstract.) The personalized videos may be made from "a portion of a current or classic movie or television show, an entire film or TV show, an advertisement, a music video, or a specialty clip made specifically for personalization (for example, a clip that can be personalized to show the new actor with a 'celebrity friend')". (*Id*. at 5:20-25.) The personalized videos are made by editing the original video (*e.g.*, the original movie or TV show) with new image information and then displaying it. (*Id*. at 2:41-54.) In other words, Senftner discloses "[a]n interactive media apparatus for generating a displayable edited video data stream from an original video data stream".

Thus, Senftner discloses the limitations of [1-PREAMBLE-i].

**[1-PREMABLE-ii] . . . wherein at least one pixel in a frame of said original video data stream is digitally extracted to form a first image, said first image then replaced by a second image resulting from a digital extraction of at least one pixel in a frame of a user input video data stream, said apparatus comprising:**

As used in the claims of the ’591 patent, the limitation “a first image” refers to an image digitally extracted (*e.g.*, copied) from a frame of the original video data stream (*e.g.*, an image or video of the original actor extracted from a frame of the original broadcast), and the limitation “a second image” refers to an image digitally extracted (*e.g.*, copied) from a frame of the user’s video data stream (*e.g.*, an image or video taken from a frame of the user’s video). Per the requirements of the preamble, the user’s image (the second image) replaces the original actor’s image (the first image) in the original broadcast (the original video data stream). This is shown in FIG. 3 of the ’591 patent, below:

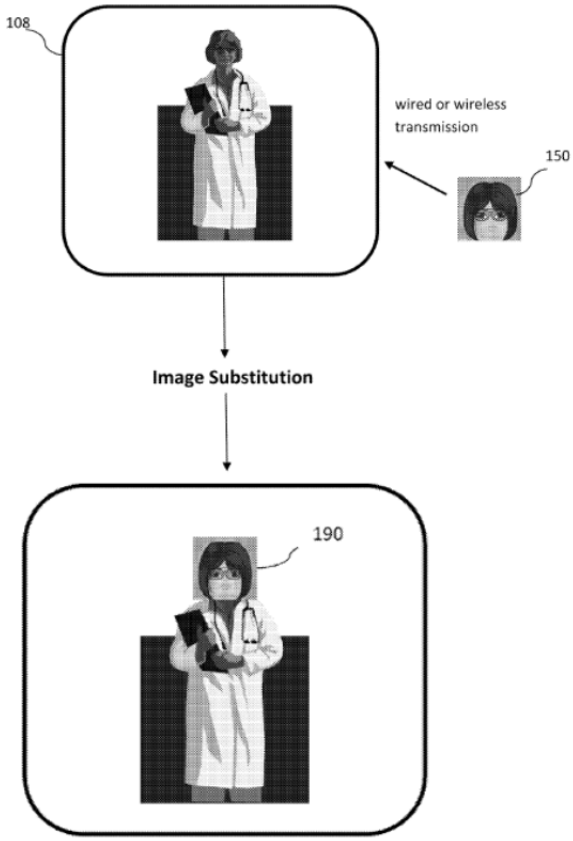



***FIG. 3***

Regarding this FIG. 3, the '591 patent states:

> "Referring now to FIG. 3 we show a user input 150 of *a photo image of the user used to replace the face of the image shown on the device 108*. The user transmits *the photo image 150* by wired or wireless means to the device 108. *The image substitution is performed and the device 108 shows the substituted image 190*." (Ex. 1001 at 2:66-3:4 (emphasis added).)

Senftner discloses this same arrangement. Senftner discloses that a "target" or "original actor" is selected in an original digital video (an "original video data

stream". (Ex. 1006 at 2:41-51.) Also, a digital video of the "target replacement" or "new actor" is captured, analyzed and stored (*Id*. at FIG. 1, 10:3-28), such as "by means of a digital image recording device 420, such as a digital camera, a digital video recorder, or a camera-equipped cell phone". (*Id.* at 17:46-49; *see also*, *id*. at 5:33-40.) Video captured with any of these devices produces "a user input video data stream".[4] Thereafter, images of the target or original actor (the first images) are "replaced" with images of a "target replacement" or "new actor" (the second images). (*Id.* at FIGS. 1-2, 2:51-54.) In one specific embodiment, an original actor's face is replaced with a new actor's face:[5]

[4] The '591 patent uses the same digital equipment disclosed by Senftner to create their "user input video data stream." (*See, e.g.*, Ex. 1001 at 2:12 ("camera-enabled personal device 106"), 2:34-35 ("a camera input"), and 3:41-43 ("the User Data Device (UDD) 106 is an image capable digital device . . .").

[5] In order to streamline the Petition, Petitioner generally refers only to "original actor" and "new actor" when explaining Senftner, as opposed to referring both "target" and "original actor," or both and "new actor" and "target replacement." However, Petitioner is not limiting the disclosure of Senftner via its streamlining.

> "The initial description of the processes will be made using an example case *where the video is personalized by substituting the image of the face of a new actor for the facial portion of the image of one of the video's original actors*." (*Id.* at 9:6-9.)

To complete this "replacement" the original video having the original actor can be altered on a "pixel-by-pixel and frame-by-frame basis". (*Id.* at 8:52-54.) Specifically, the "replacement" process may involve "overwriting the original data with the new data" in a single step. (*Id.* at 8:67-9:1.) In this regard:

> "The personalization process begins at step 320 *where the image of the new actor is inserted into the video*. The process for substituting the image of the new actor is show [*sic*] in additional detail in FIG. 2. At step 322, *the 3D model of the new actor may be transformed to match the orientation and expression of the original actor as defined by data from step 210 of the video preparation process*. This transformation may involve both rotation on several axis and geometric morphing of the facial expression, in either order. After the 3D model is rotated and morphed, a 2D image of the 3D model is developed and scaled to the appropriate size at step 324. *The*

---

Senftner refers to other targets and target replacements besides actors. (*See, e.g.*, Ex. 1006 at 2:44-46, 13:15-25.)

> *transformed scaled 2D image of the new actor is then inserted into the video at step 326 such that the position, orientation, and expression of the new actor substantially matches the position, orientation, and expression of the previously removed original actor*. In this context, a "substantial match" occurs when the personalized video presents a convincing illusion that the new actor was actually present when the video was created." (*Id.* at 12:27-45.)

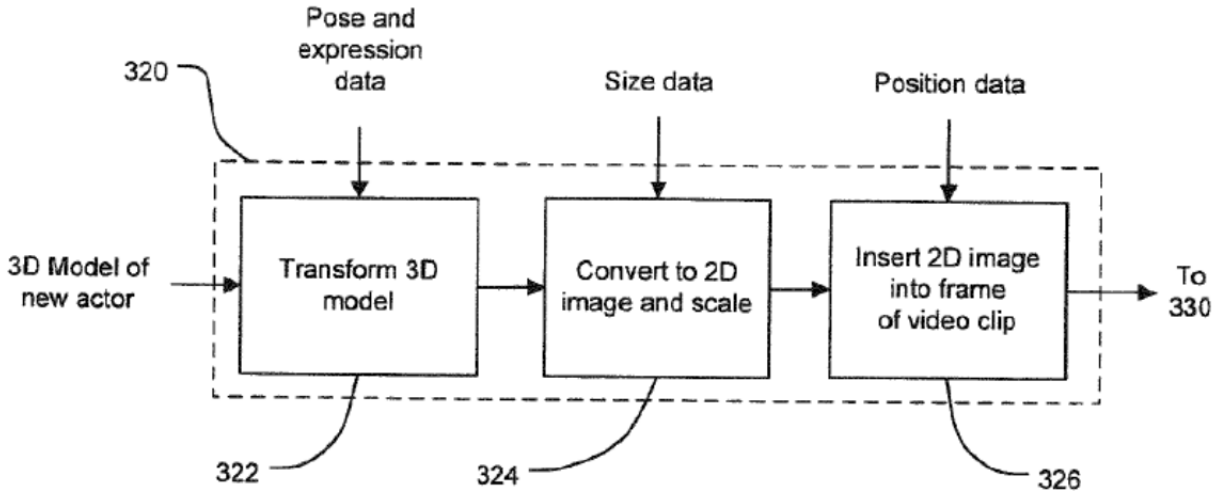



FIG. 2

Senftner expressly discloses that the original actor is removed from the original digital video:

> "In order for such alteration to occur, the replacement of the face and portions of the head is not enough to achieve this result; in this situation *a complete removal of the original actor is executed*, their key motions are preserved in a secondary storage medium, and then referenced for the animation and insertion of the petite female's digital double." (*Id.* at 6:8-14.)

(*See also, id.* at 5:15-25, 11:7-12, 42-59.) The replacement of the original actor by the new actor may be completed for a single image or for a series of images. (*Id.* at 5:19-28.)

Thus, Senftner discloses the limitation "wherein at least one pixel in a frame of said original video data stream is digitally extracted to form a first image, said first image then replaced by a second image resulting from a digital extraction of at least one pixel in a frame of a user input video data stream," and therefore discloses the limitations of [1-PREAMBLE-ii].

**[1a] an image capture device capturing the user input video data stream;**

The '591 patent discloses various "image capture devices" for capturing the "user input video data stream". (Ex. 1001 at 2:12, FIG. 1.) Senftner also discloses "image capture devices" and "user input video streams": "[t]he 2D digital image 425 may be created by means of a digital image recording device 420, such as a digital camera, a digital video recorder, or a camera-equipped cell phone." (Ex. 1006 at FIGS. 8-11, 17:45-48.) The disclosed cameras capture "user input video data streams". Thus, Senftner discloses limitation [1a].

**[1b] an image display device displaying the original video stream;**

The '591 patent discloses various display devices for displaying the original video stream, such as "a digital TV, a gaming console, a digital media

player/recorder, or set top box . . . ." (Ex. 1001 at 2:37-39.) Senftner also discloses display devices: "[t]he personalized video may then be presented to user 650 by means of *display device*, and may be stored in memory 720 or storage medium 730". (Ex. 1006 at 21:6-8.) FIG. 10 also shows a computer (670) with a monitor that may be used to display the original video data stream. Thus, Senftner discloses limitation [1b].

**[1c-i] a data entry device, operably coupled with the image capture device and the image display device**

Senftner discloses data entry devices meeting the requirements of limitation [1c-i]. Figure 10 of Senftner shows a computer (670) having a keyboard (not numbered) and a display device (a monitor) for "for creating personalized videos." (*Id.* at 20:24.) As shown in FIG. 10, the computer is operably coupled to digital image sources (660) (*i.e.*, one or more image capture devices), such as any of the digital camera, cell phone, and digital video recorder illustrated in FIG. 10. (*Id.* at 20:35-36.) A similar arrangement is show in FIG. 11, in which the computing device (700) includes an interface to requestor (650) "such as a keyboard, mouse, or other human interface means" and "may also have an interface to a digital image device 660" which is the image capture device. After personalizing the digital video, the personalized video "may then be presented to user 650 by means of display device. . . ." (*Id.* at 20:62-64, 21:5-7.)

Thus, Senftner discloses limitation [1c-i].

**[1c-ii] . . . operated by a user to select the at least one pixel in the frame of the user input video data stream to use as the second image, and further operated by the user to select the at least one pixel to use as the first image;**

Building on the discussion of limitation [1c-i], Senftner discloses that the computer's user (the "requestor" (650)) operates the computer through the data entry device. (*Id.* at 18:1-18, 20:35-38.) This "requestor" is shown in FIGS. 8-11, and meets the requirements of "a user" who selects the first and second images, as required by limitation [1c-ii]. FIGS. 8 and 9 show a flow chart of a process 400/500 for creating and delivering a personalized video. (*Id.* at 17:23-24, 18:45-46.) FIGS. 10 and 11 describe the computing devices for creating the personalized video. (*Id.* at 5:5-6, 20:24-25, 20:56-57.)

In replacing the original actor's image with the new actor's image, the requestor "selects" both the original actor's and new actor's images:

> "Apparatus, systems and techniques for providing personalized digital video in various applications are described. *One or more target images, such as an actor and an object, in an original digital video can be replaced based on user preferences to produce a personalized digital video*. . . . In one implementation, a computer-implemented process for providing personalized digital video can include *selecting a target in original digital video* to be replaced by a target replacement, wherein the target is a portion or an entirety of an actor or an object

> other than an actor in the original digital video . . ." (*Id.* at 2:33-45 (emphasis added).)
>
> "The act of 'replacing' may involve *identifying all pixels within each video frame that represent an image of the original object to be replaced* . . ." (*Id.* at 8:60-62 (emphasis added).)
>
> "*The transformed scaled 2D image of the new actor is then inserted into the video at step 326 such that the position, orientation, and expression of the new actor substantially matches the position, orientation, and expression of the previously removed original actor*. In this context, a "substantial match" occurs when the personalized video presents a convincing illusion that the new actor was actually present when the video was created." (*Id.* at 12:37-45.)

Senftner acknowledges and a POSITA would understand that each frame of a digital video is comprised of pixels. (*Id.* at 2:8-10) (Ex. 1003 at ¶41.) Senftner further discloses that each frame of a video contains a 2D image. (Ex. 1006 at 11:57.) Thus, a POSITA would understand that, when a 2D image (a second image), primarily capturing the new actor's face, is selected by the requester for the actor modeling process, the requester also necessarily selects the at least one pixel comprising the selected 2D image. (*Id.* at 2:41-45, 10:3-16, Fig. 10, 20:24-39.) (Ex. 1003 at ¶¶42-43.) Likewise, a POSITA would understand that when the requester selects an original object (target/first image) to be replaced by a target replacement (second image), the requester necessarily selects at least one pixel

comprising the image of the original object. (Ex. 1006 at 2:41-45, 8:60-64, Fig. 10, 20:24-39.) (Ex. 1003 at ¶¶42, 44.) Indeed, Senftner discloses that the disclosed replacement process manipulates a digital video (original video data stream) on a pixel-by-pixel and frame-by-frame basis, indicating that the pixel information of the first image and the second image must be used by the system. (Ex. 1006 at 8:52-67) (Ex. 1003 at ¶42.)

Thus, Senftner discloses limitation [1c-ii].

**[1d] wherein said data entry device is selected from a group of devices consisting of: a keyboard, a display, a wireless communication capability device, and an external memory device;**

As shown in limitation [1c-i], the computer of Senftner may use a keyboard. (Ex. 1006 at 20:65-67.) In addition, Senftner discloses other types of data entry devices such as an external memory device, a display, and an Internet capability device. (*Id.* at 20:67-21:11.) Thus, Senftner discloses limitation [1d].

**[1e] a digital processing unit operably coupled with the data entry device, said digital processing unit performing:**

"Digital processing unit" is not defined in the '591 patent. "Digital processing unit" also appears to have no special meaning in the '591 patent. Therefore, any computer capable of performing the functions claimed below meets the "digital processing unit" requirements of claim 1.

The computers of Senftner include normal hardware (*e.g.*, a processor 610/710) and software. (*Id.* at FIGS. 10-11 and 20:56-21:11.) The computers of Senftner also perform the claimed functions required by limitations [1e-i] through [1e-vii], below. Thus, Senftner discloses limitation [1e].

**[1e-i] identifying the selected at least one pixel in the frame of the user input video data stream;**

As explained above relative to limitation [1c-ii], Senftner discloses that the requester (user) selects the new actor's image. (*Id.* at 10:3-12, 12:27-45, 18:1-22.) Senftner further discloses matching pixel(s) from the original image to pixel(s) in the new image, on "a pixel-by-pixel and frame-by-frame basis" to achieve the replacement. (*Id.* at 8:52-9:5.) Senftner further discloses that portions of the original image in the frame of a video must be identified to achieve the replacement. (*Id.* at 8:60-62.)

Thus, Senftner discloses limitation [1e-i].

**[1e-ii] extracting the identified at least one pixel as the second image;**

As discussed in [1c-ii] and [1e-i], above, Senftner's system performs the replacement process on a pixel-by-pixel and frame-by-frame basis, and a POSITA would understand that selecting/extracting an image necessarily requires selecting/extracting the pixel information relating to the selected image. (Ex. 1003

at ¶49.) Senftner discloses that, after the selection, the "new actor" image (the second image) may be identified and extracted (copied):

> "*The actor modeling process 100 accepts one or more two-dimensional (2D) digital images of the new actor*, plus related supporting information, and creates, at step 110, a digital model of the new actor composed of a three-dimensional model and, optionally, a demographic profile and other personal information describing the new actor. *The preferred 2D image primarily captures the new actor's face, the top and bottom of their head, both ears, portions of their neck, with both eyes visible and no more than a 30 degree rotation away from the camera*." (Ex. 1006 at 10:3-12.)
>
> "In another implementation, a process for personalizing a video can include providing a video library of a plurality of prepared videos, each of the prepared videos resulting from a video preparation process; *providing an actor model library of one or more new actor models where each of the models resulting from an actor modeling process*; selecting a video from the video library; *selecting a new actor model from the actor model library*; and applying a personalization process to create a personalized version of the selected video using the selected new actor model." (*Id.* at 4:15-24.)

Since images of the new actors are captured in an actor modeling process, and stored in an actor digital library, Senftner discloses "extracting the identified at least one pixel as the second image," and thus discloses limitation [1e-ii].

Nonetheless, in the event Patent Owner argues a narrow construction of "digital extraction" in that pixel(s) must actually be removed from the data stream,[6] and the Board accepts this construction, Senftner still renders obvious this limitation. As shown above, Senftner discloses copying the applicable pixel(s) of the new actor's image during the new actor modeling process and that the original actor's image (first image) in the original video file "is removed" (*see, e.g.,* 2:51-54, 58-62). Thus, a POSITA would find it obvious to also remove the new actor's image (second image) from its applicable data stream. Indeed, a POSITA would find this to be just one of two ways to obtain new actor images (copying the data or removing or cutting the data), and it would be simple and routine to change from a "copying" process, where pixels in the original are maintained, to a "cutting" process, where pixels in the original are removed. (Ex. 1003 at ¶51.) This is a simple matter of deleting the information from the original data stream when it is copied for use in the substitution processing. (*Id.*) Thus, even if "digital

[6] Such a narrow construction would not comport with the written description requirement of §112, as Patent Owner does not show possession of removal of pixels in the specification of the '591 patent.

extraction" is unduly limited to removal of the data from the user's data stream, Senftner still teaches or suggests the limitation of "extracting the identified at least one pixel as the second image," and it would be a simple and routine modification to cut (remove) the applicable pixel data from the applicable data stream, as opposed to copying those pixels, as disclosed by Senftner. (*Id.*)

**[1e-iii] storing the second image in a memory device operably coupled with the interactive media apparatus;**

Senftner discloses limitation [1e-iii]. FIG. 8 shows that the "new actor" images from step 100 are stored in the "Actor Model Library" (440). (Ex. 1006 at 17:65-67.) This library is coupled to the computer in order to achieve the replacement. (*Id.* at 18:11-12.) Moreover, Senftner discloses a memory device operably coupled with the disclosed system. (*Id.* at 21:23-29.) Thus, Senftner discloses limitation [1e-iii].

**[1e-iv] receiving a selection of the first image from the original video data stream;**

As described above, Senftner discloses that the computer receives a selection of the original actor's face (the first image) from the original video (original video data stream) in order to replace the original actor's face (first image) with the new actor's face (second image). (*Id.* at 2:41-44, 9:6-9.) This process is also more particularly described relative to FIGS. 1-2 and process 200 of Senftner (*Id.* at 10:29-12:17.)

Thus, Senftner discloses limitation [1e-iv].

**[1e-v] extracting the first image;**

As disclosed above relative to claim limitations [1-PREAMBLE-ii] and [1e-iv], the first image (*e.g.*, the original actor's face) is selected and separated out (extracted) for replacement by the second image (*e.g.*, the "new actor's" face). (*See, e.g.,* Ex. 1006 at 2:33-54, 5:42-59, 6:8-14, 8:58-9:5, and 11:7-12.)

Thus, Senftner discloses limitation [1e-v].

**[1e-vi] spatially matching an area of the second image to an area of the first image in the original video data stream, wherein spatially matching the areas results in equal spatial lengths and widths between said two spatially matched areas; and**

Senftner discloses limitation [1e-vi]. In order to replace the original actor's image (first image) with the new actor's image (second image), the images must necessarily be spatially matched in the X-Y dimensions (length-width). (*Id.* at 10:29-46, 12:27-45.) (Ex. 1003 at ¶55.) Thus, Senftner discloses limitation [1e-vi].

**[1e-vii] performing a substitution of the spatially matched first image with the spatially matched second image to generate the displayable edited video data stream from the original video data stream**

As explained in limitations [1e-iv] and [1e-ii], above, the new actor image replaces the original actor video. (Ex. 1006 at 2:33-54; *see also, e.g., id. at* 5:42-59, 8:58-9:5, and 12:27-45.) Thus, Senftner discloses limitation [1e-vii].

For these reasons, Senftner discloses all limitations of claim 1, and anticipates claim 1. For also these reasons, Senftner discloses, teaches or suggests all limitations of Claim 1, rendering Claim 1 obvious.

**[Claim 2] The interactive media apparatus of claim 1 wherein the digital processing unit is further capable of performing: computing motion vectors associated with the first image; and applying the motion vectors to the second image extracted from the user input video data stream, wherein the generated displayable edited video data stream resulting from the substitution maintains an overall motion of the original video data stream.**

Regarding claim 2, the '591 patent states:

> "The process further requires the steps of background analysis, face detection and recognition, image and video motion detection and analysis to follow through the input device's motion sequence with the user's inputted data to produce a resulting stream where the user inputted data becomes part of the input device's stream as if it were originally its own." (Ex. 1001 at 3:28-40.)
>
> "To accomplish the embedding process, the internet network devices are capable of performing at least the following functions: . . . *motion detection and estimation, motion correction to adapt the motion of the user input sequence with that of the original metadata to be broadcasted*, error analysis, etc." (*Id.* at 4:28-40.)

(*See also*, *id.* at 4:57-5:4 and 5:38-59.)

Senftner discloses the same motion detection and analysis for both the first and second image in order to substitute the new actor for the original actor. As

shown below, the original actor's "motion vectors" are captured to facilitate their replacement with the new actor's motion vectors:

> "In one implementation, a *computer-implemented* process for providing personalized digital video can include . . . *analyzing each frame of the original digital video to track a change in the selected target in the original digital video* to capture data on the selected target, wherein the captured data includes at least information on a *position, orientation and size of the selected target in the original digital video*; and *replacing the selected target with an image* that resembles a continuation of a scene adjacent to the target in the original digital video to produce altered digital video in which the selected target is removed." (Ex. 1006 at 2:41-54.)
>
> "In order for such alteration to occur, the replacement of the face and portions of the head is not enough to achieve this result; in this situation a complete removal of the original actor is executed, *their key motions are preserved in a secondary storage medium, and then referenced for the animation and insertion of the petite female's digital double*." (*Id.* at 6:8-14.)

Thus, Senftner discloses "wherein the digital processing unit is further capable of performing: computing motion vectors associated with the first image".

Senftner also discloses that the new actor's "motion vectors" may be captured so that the new actor can be substituted for the original actor in the original video. (*Id.* at 17:10-23.) Hence, Senftner discloses "applying the motion

vectors to the second image extracted from the user input video data stream, wherein the generated displayable edited video data stream resulting from the substitution maintains an overall motion of the original video data stream."

Thus, Senftner discloses and anticipates claim 2.

**[Claim 8] The interactive media apparatus of claim 1, wherein the substitution performed by the digital processing device replaces at least a face of a first person from the original video data stream by at least a face of a second person from the user input video data stream.**

As discussed above relative to claim 1, Senftner discloses replacing an original actor's face with a new actor's face. (*Id.* at 9:6-9.)

Thus, Senftner discloses and anticipates claim 8.

**[Claim 11]**

Petitioner provides the following comparison chart of the limitations of claim 11 compared with the equivalent limitations of claims 1 and 2. As the chart shows, these limitations were addressed previously with respect to claims 1 and 2, and thus the limitation-by-limitation analysis for claim 11 identifies the corresponding limitation that was previously addressed, rather than repeat the same analysis.

| | |
|---|---|
| [11P-i] A method for generating a displayable edited video data stream from an original video data stream, | [1P-i] An interactive media apparatus for generating a displayable edited video data stream from an original video data stream, |

| | |
|---|---|
| [11P-ii] wherein at least one pixel in a frame of the original video data stream is digitally extracted to form a first image, said first image then replaced by a second image resulting from a digital extraction of at least one pixel in a frame of a user input video data stream, said method comprising: | [1P-ii] wherein at least one pixel in a frame of said original video data stream is digitally extracted to form a first image, said first image then replaced by a second image resulting from a digital extraction of at least one pixel in a frame of a user input video data stream, said apparatus comprising: |
| [11a] capturing a user input video data stream by using a digital video capture device; | [1a] an image capture device capturing the user input video data stream; |
| | [1b] an image display device displaying the original video stream; |
| [11b] using a data entry device operably coupled with the digital video capture device and a digital display device, selecting the at least one pixel in the frame of the input video data stream; | [1c] a data entry device, operably coupled with the image capture device and the image display device, operated by a user to select the at least one pixel in the frame of the user input video data stream to use as the second image, and further operated by the user to select the at least one pixel to use as the first image; |
| [11c] wherein the data entry device is selected from a group of devices consisting of: a keyboard, a display, a wireless communication capability | [1d] wherein said data entry device is selected from a group of devices consisting of: a keyboard, a display, a wireless communication capability |

| | |
|---|---|
| device, and an external memory device; and | device, and an external memory device; |
| [11d] using a digital processing unit operably coupled with the data entry device, performing: | [1e] a digital processing unit operably coupled with the data entry device, said digital processing unit performing: |
| [11d-i] identifying the selected at least one pixel in the frame of the input video stream; | [1e-i] identifying the selected at least one pixel in the frame of the user input video data stream; |
| [11d-ii] extracting the identified at least one pixel as the second image; | [1e-ii] extracting the identified at least one pixel as the second image; |
| [11d-iii] storing the second image in a memory device operably coupled with the digital processing unit; | [1e-iii] storing the second image in a memory device operably coupled with the interactive media apparatus; |
| [11d-iv] receiving a selection of the first image from the user operating the data entry device; | [1e-iv] receiving a selection of the first image from the original video data stream; |
| [11d-v] extracting the first image from the original video data stream; | [1e-v] extracting the first image; |
| [11d-vi] spatially matching an area of the second image to an area of the first image in the original video data stream, wherein spatially matching the areas results in equal spatial lengths and widths between said two spatially matched areas; | [1e-vi] spatially matching an area of the second image to an area of the first image in the original video data stream, wherein spatially matching the areas results in equal spatial lengths and widths between said two spatially matched areas; and |

| | |
|---|---|
| [11d-vii] performing a substitution of the spatially matched first image with the spatially matched second image to generate a the displayable edited video data stream from the original video data stream; | [1e-vii] performing a substitution of the spatially matched first image with the spatially matched second image to generate the displayable edited video data stream from the original video data stream. |
| [11d-viii] computing motion vectors associated with the first image; and | Claim 2, limitation [2a] The interactive media apparatus of claim 1 wherein the digital processing unit is further capable of performing: computing motion vectors associated with the first image; and |
| [11d-ix] applying the motion vectors to the second image, wherein the generated displayable edited video data stream resulting from the substitution maintains an overall motion of the original video data stream. | Claim 2, limitation [2b] applying the motion vectors to the second image extracted from the user input video data stream, wherein the generated displayable edited video data stream resulting from the substitution maintains an overall motion of the original video data stream. |

**[11P-i] A method for generating a displayable edited video data stream from an original video data stream,**

Senftner discloses the limitations of [11P-i] for the reasons described above with respect to [1-PREAMBLE-i].

**[11P-ii] wherein at least one pixel in a frame of the original video data stream is digitally extracted to form a first image, said first image then replaced by a second image resulting from a digital extraction of at least one pixel in a frame of a user input video data stream, said method comprising:**

Senftner discloses the limitation of [11P-ii] for the reasons described above with respect to [1-PREAMBLE-ii].

**[11a] capturing a user input video data stream by using a digital video capture device;**

Senftner discloses limitation [11a] for the reasons described above with respect to limitation [1a].

**[11b] using a data entry device operably coupled with the digital video capture device and a digital display device, selecting the at least one pixel in the frame of the input video data stream;**

Senftner discloses limitation [11b] for the reasons described above with respect to limitation [1c].

**[11c] wherein the data entry device is selected from a group of devices consisting of: a keyboard, a display, a wireless communication capability device, and an external memory device; and**

Senftner discloses limitation [11c] for the reasons described above with respect to limitation [1d].

**[11d] using a digital processing unit operably coupled with the data entry device, performing:**

Senftner discloses limitation [11d] for the reasons described above with respect to limitation [1e].

**[11d-i] identifying the selected at least one pixel in the frame of the input video stream;**

Senftner discloses limitation [11d-i] for the reasons described above with respect to limitation [1e-i].

**[11d-ii] extracting the identified at least one pixel as the second image;**

Senftner discloses limitation [11d-ii] for the reasons described above with respect to limitation [1e-ii].

**[11d-iii] storing the second image in a memory device operably coupled with the digital processing unit;**

Senftner discloses limitation [11d-iii] for the reasons described above with respect to limitation [1e-iii].

**[11d-iv] receiving a selection of the first image from the user operating the data entry device;**

Senftner discloses limitation [11d-iv] for the reasons described above with respect to limitation [1e-iv].

**[11d-v] extracting the first image from the original video data stream;**

Senftner discloses limitation [11d-v] for the reasons described above with respect to limitation [1e-v].

**[11d-vi] spatially matching an area of the second image to an area of the first image in the original video data stream, wherein spatially matching the areas results in equal spatial lengths and widths between said two spatially matched areas;**

Senftner discloses limitation [11d-vi] for the reasons described above with respect to limitation [1e-vi].

**[11d-vii] performing a substitution of the spatially matched first image with the spatially matched second image to generate a the displayable edited video data stream from the original video data stream;**

Senftner discloses limitation [11d-vii] for the reasons described above with respect to limitation [1e-vii].

**[11d-viii] computing motion vectors associated with the first image; and**

Senftner discloses limitation [11d-viii] for the reasons described above with respect to claim 2, limitation [2a].

**[11d-ix] applying the motion vectors to the second image, wherein the generated displayable edited video data stream resulting from the substitution maintains an overall motion of the original video data stream.**

Senftner discloses limitation [11d-ix] for the reasons described above with respect to claim 2, limitation [2b].

Thus, Senftner anticipates claim 11.

### B. Ground 2: Claims 3 and 4 are obvious over Senftner in view of Levoy[7]

#### 1. Scope and Content of the Prior Art

For Ground 2 the prior art consists of Senftner and Levoy. Senftner is described in detail above in Ground 1, and is pre-AIA 102(b) prior art to the '591 patent.

Levoy (US2009/0309990) was filed on June 11, 2008, and published on December 17, 2009. Levoy is therefore pre-AIA §§102(a),(e) prior art to the '591 patent.

Levoy describes systems and methods for selecting one or more portions of an image via a touch screen device, and then incorporating those selected portions into another image to creating a new composite image, such as on a mobile phone or other touch screen device. As one example, FIG. 3 of Levoy (reproduced below) illustrates how an image fragment (330, 340 or 350), displayed as a list on

[7] As shown herein, claims 3-4 contain numerous §112 issues (*e.g.*, written description, indefiniteness). Nonetheless, solely for purposes of this Petition, Petitioner construes these claims to show their invalidity. Petitioner's constructions of these claims is not a concession that the claims are complaint with §112.

the screen, may be incorporated into an underlying image (310) using a touch screen device.

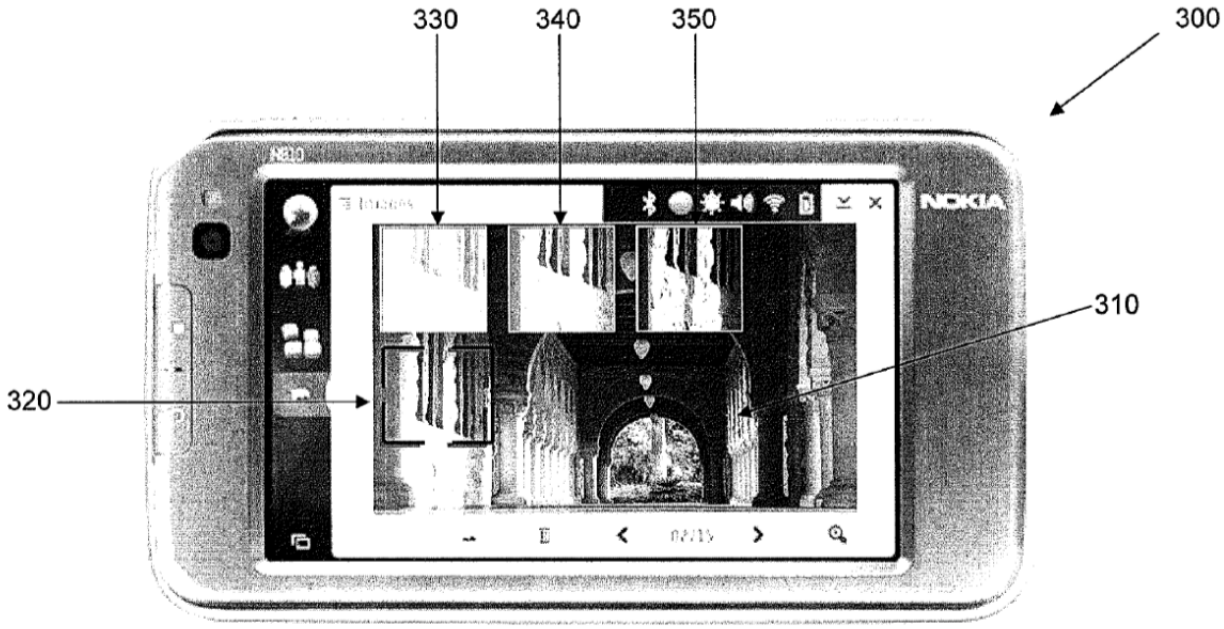



FIG. 3

Specifically, a sample burst image 310 is presented on the display of the mobile device 300. (Ex. 1008 at ¶46.) The mobile device 300 may be a touch screen device. (*Id*. at, *inter alia*, ¶¶21, 23, and, 47.) When a user selects boxed area 320, such as via touching boxed area 320 with a fingertip, image fragments 330, 340, 350 are generated. (*Id*. at ¶¶46-47.) A user may then select one or more of the image fragments for incorporation into the underlying image. (*Id*. at ¶¶46, 50.) Regarding image selection and incorporation, Levoy explains:

"[T]he apparatus 100 may include various means *for receiving a selection of a particular burst image*, which may include the processor 105, the presenter 134, the user interface 115, *a display* (e.g., *a touch screen display* or a conventional display), algorithms executed by the foregoing or other elements for receiving a selection of a particular burst image described herein and/or the like. In this regard, *a user may interact with the user interface 115 to select one of the burst images via the presentations of the burst image fragments*. For example, *a user may tap on a touch screen in the location of a particular burst image fragment to select the underlying burst image*. The selection may be obtained by the user interface 115 and transmitted to the processor 105 to be received by the processor 105. (*Id*. at ¶47.)

"*The processor 105 may also be configured to generate* ***a composite image*** *based on one or more selected burst images and the corresponding one or more selected locations associated with the selected burst images*. In some exemplary embodiments, the processor may also be configured to *provide for the presentation of the composite image after generation*. In this regard, a composite image may be generated in any known manner. However, the inputs to the generation of the composite image may be derived for the selected burst images and the selected locations associated with the selected burst images." (*Id*. at ¶50.)

**2. Motivation to Combine**

As discussed above in [1-PREAMBLE-i], Senftner discloses systems and methods for personalizing video through partial image replacement. (Ex. 1006 at Abstract, 2:41-54.) Likewise, Levoy discloses systems and methods for image editing that creates a composite image based on the selection of a portion of a sample burst image and the selection of a corresponding portion of another burst image. (Ex. 1008 at Abstract, ¶¶58-61.) Accordingly, both Senftner and Levoy are directed to the same fundamental process of substituting a portion of an image data with a portion of another image data, creating a new image.

Senftner discloses various computer devices that can be used to select the original images and new images for replacement, such as:

> "personal computers, server computers, *computing tablets*, set top boxes, video game systems, personal video recorders, telephones, *personal digital assistants (PDAs), portable computers, and laptop computers*." (Ex. 1006 at 21:25-29.)

Senftner does not specifically disclose that any of these devices inherently has a touch screen.

However, as evidenced by Levoy, touch screen devices were generally well-known at the time of the '591 patent was filed. (Ex. 1008 at, *inter alia*, ¶¶10, 21.) Indeed, well before the filing of the '591 Patent, touch screen devices were known

to be more beneficial than keyboards for their "ease and versatility of operation" (Ex. 1010 at ¶4, *see also,* Ex. 1011 at 1:25-36), and known to be "the most natural and user friendly to the operator". (Ex. 1012 at 1:24-40 ("Human factors studies have shown that by providing a means for inputting data on the visual display screen itself, the user can achieve the most closely coupled interactive operations with the data processing system. When the user responds to visual signals output at the face of the visual display device, by inputting signals at that same visual display surface, an accuracy and immediacy in the interaction between man and machine can be achieved. This form of input device is easy to learn to use and seems the most natural and user-friendly to the operator.").) (Ex. 1003 at ¶81.) Thus, instead of using a conventional non-touch screen computer as the computerized device of Senftner, a POSITA would find it obvious to, instead, use conventional touch screen technology with any one of the computing devices enumerated by Senftner, including, for instance, the "computing tablets . . . personal digital assistants (PDAs), portable computers, and laptop computers" recited by Senftner. (*Id.* at ¶82.)

### 3. Claims

**[Claim 3] The interactive media apparatus of claim 1 wherein the digital processing unit is further capable of extracting the at least one pixel from the user entering data in the data entry display device.**

Petitioner first notes that Claim 3 is invalid on multiple §112 grounds, including written description and indefiniteness. The specification of the '591 patent does not disclose a digital processor capable of extracting pixel(s) "from the user entering data in the data entry display device." Thus, claim 3 is invalid for lack of written description.

Claim 3 is also indefinite. There is no antecedent basis for the term "the data entry display device". Claim 1 recites a "data entry device," separate "image display device" and recites "wherein the data entry device is selected from the group consisting of . . . a display". No "data entry display device" is recited. Nonetheless, for the purposes of this IPR only, Petitioner assumes that "data display entry device" of claim 3 means the "data entry device" of claim 1 is a display, as per the Markush group of claim 1.

Claim 3 is also indefinite because it is not clear whether "at least one pixel" refers to the "at least one pixel" of the "original video data stream" or the "user input data stream." Nonetheless, for purposes of this IPR only, Petitioner shows that claim 3 is invalid when the "at least one pixel" of claim 3 is construed to refer to the "user input data stream" of claim 1.

Assuming claim 3 can be construed, Senftner in combination with Levoy renders claim 3 obvious. As shown above relative to claim 1, Senftner discloses "the digital processing unit is further capable of extracting the at least one pixel from the user entering . . ." limitation of claim 3. Claim 3 further limits claim 1 by requiring that the user enter the data "in the data display entry device". As noted above, Petitioner assumes "the data display entry device" means the "data entry device" is a display. In other words, "the data display entry device" is a touch screen or other similar display device capable of data entry.

Senftner discloses various computer devices that can be used to select the original images and new images for replacement, such as:

> "personal computers, server computers, *computing tablets*, set top boxes, video game systems, personal video recorders, telephones, *personal digital assistants (PDAs), portable computers, and laptop computers*." (Ex. 1006 at 21:25-29.)

Senftner does not specifically disclose that any of these devices inherently has a touch screen.

However, as described in the Motivation to Combine section (VII.B.2) above, and as evidenced by Levoy, touch screen devices were generally well-known at the time of the '591 patent was filed, and were known to provide benefits over non-touch screen devices. Thus, instead of using a conventional non-touch

screen computer as the computerized device of Senftner, a POSITA would find it obvious to use conventional touch screen technology with any one of the computing devices enumerated by Senftner, including, for instance, the "computing tablets . . . personal digital assistants (PDAs), portable computers, and laptop computers" recited by Senftner, which would allow a user to copy (extract) pixels from the user entering data in the touch screen (data entry display device). (Ex. 1003 at ¶¶81-82.) (§VII.B.2, above.) Accordingly, Senftner, modified to have a touch screen per Levoy, discloses the limitations of claim 3, rendering claim 3 obvious.

**[Claim 4] The interactive media apparatus of claim 3 wherein the digital processing unit is further capable of extracting the at least one pixel from the user pointing to a spatial location in a displayed video frame,**

Like claim 3, claim 4 is invalid under §112 for lack of written description and for being indefinite. The specification of the '591 patent does not show possession of a digital processor capable of extracting pixel(s) "from the user pointing to a spatial location in a displayed video frame" and is therefore invalid for lack of a written description.

Claim 4 is also indefinite due to the same "at least one pixel" issue described above for claim 3. Nonetheless, for purposes of this IPR only, Petitioner shows that claim 4 is invalid when the "at least one pixel" of claim 4 is construed to refer to the "user input data stream" of claim 1.

Senftner in view of Levoy teaches or suggests the limitations of claim 4. As described above relative to claim 3, Senftner does not specifically disclose using a touch screen device as the computing device. As also described above relative to claim 3, it would be obvious in view of Levoy to use a touch screen computing device as the computing device of Senftner. When using touch screen technology to edit pictures, a user would naturally point "to a spatial location in a displayed video frame", as evidenced by Levoy. (Ex. 1008 at ¶47.)

Thus, Senftner modified to have a touch screen per Levoy, discloses the limitations of claim 4, rendering claim 4 obvious.

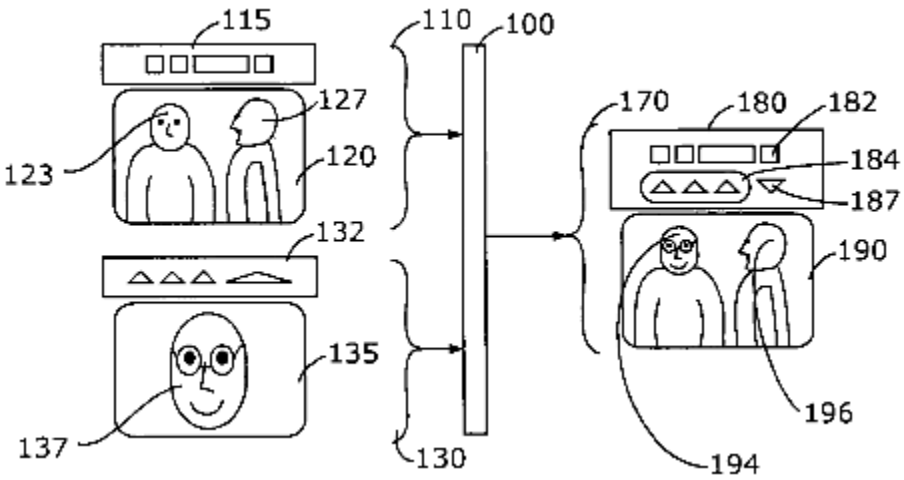
### C. Ground 3: Claims 1, 2, 8, and 11 are obvious under Sitrick

#### 1. Overview of Sitrick

Sitrick relates to "a system and method for processing a video input signal providing for tracking a selected portion in a predefined audio-visual system and integrating selected user images into the selected portion of the predefined audiovisual presentation." (Ex. 1007 at Abstract.) Like the '591 patent, Sitrick discloses computerized systems and methods in which "a user selected image [a second image] is selectively integrated into a predefined presentation in place of a tracked portion [a first image] of the predefined audiovisual presentation [an original video data stream]." (*Id.* at ¶11.)

Sitrick's Fig. 1, reproduced below, is a "system block diagram of the present invention" and provides a high-level overview of the image substitution process, substituting a facial image from an external source of user image content into an original video (program video 120) to create an edited video (output video 190).

Fig. 1




(*Id.* at Fig. 1.)

Sitrick provides that the user's image (137) is captured and provided to a subsystem 100 via user image content 130. A program video 120 is also provided to subsystem 100 via program content 110. These two contents (110 and 130) are merged via the subsystem 100, after which output content 170 is provided to a display device to show output video 190. (*Id.* at ¶31.)

> "The output video 190 consists of a processed version of the program video 120 selectively processed by the subsystem 100 *such that the representation 123 has been replaced by the user specified image 137 producing the output 194*. The input image 127 is unmodified by the

> system and output as representation 196 in the output video 190." (*Id*.)

Hence, Sitrick discloses a user input second image (face 137) is substituted for the first image (123) into the original video data stream (120) to create an edited displayable video (190).

For the reasons provided below, Sitrick discloses, teaches or suggests all limitations of claims 1, 2, 8 and 11 of the Challenged Claims, thereby rendering these claims obvious.

### 2. Claims

**[1-PREAMBLE-i] An interactive media apparatus for generating a displayable edited video data stream from an original video data stream,**

Sitrick is an interactive system that generates a displayable edited video data stream from an original video data stream. (*Id.* at Abstract.) Sitrick describes a "user image video processing and integration subsystem". (*Id*. at ¶31.) This integration subsystem 100 is "an interactive media apparatus for generating a displayable edited video data stream from an original video data stream" that receives program content 110 and user image content 130. (*Id*.) The program content 110 includes a program video 120 (an original video data stream). Subsystem 100 processes the program content 110 and user image content 130 to produce output content 170, which is displayed on a display device as output video

190. (*Id.*) The output video 190 shows the program video 120, altered "such that the representation 123 has been replaced by the user specified image 137 producing the output 194." (*Id.*)

Thus, Sitrick discloses an "interactive media apparatus for generating a displayable edited video data stream from an original video data stream," and therefore discloses the limitations of [1-PREAMBLE-i].

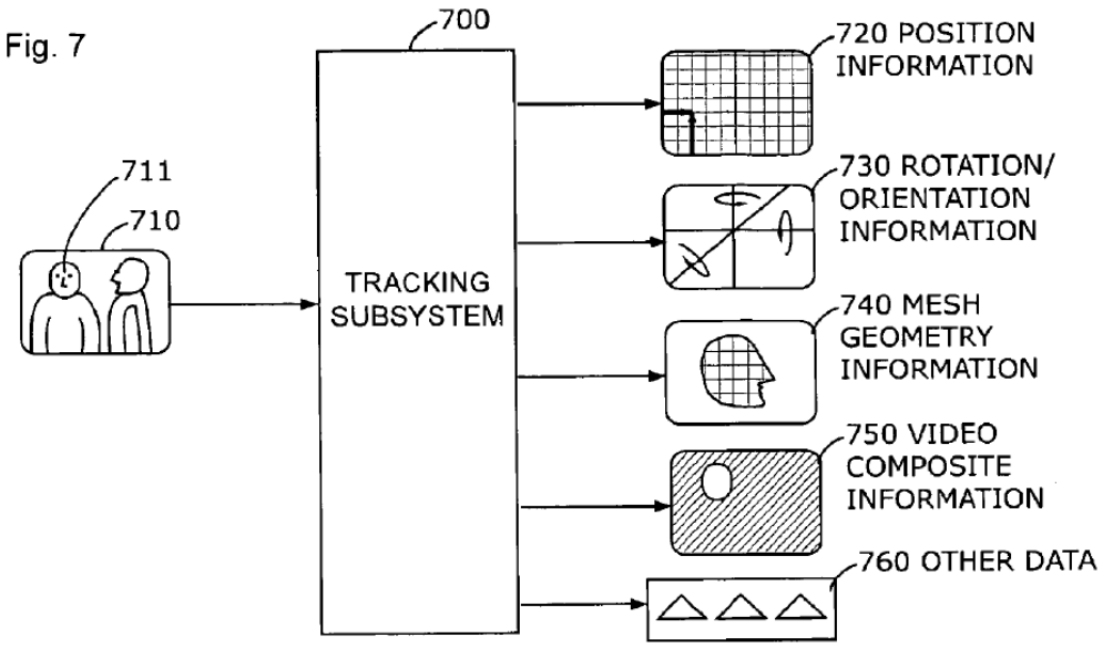
**[1-PREAMBLE-ii] . . . wherein at least one pixel in a frame of said original video data stream is digitally extracted to form a first image, said first image then replaced by a second image resulting from a digital extraction of at least one pixel in a frame of a user input video data stream, said apparatus comprising:**

Sitrick discloses the limitations of [1-PREAMBLE-ii]. Specifically, Sitrick discloses the production of a "mask" (a first image) from pixel data digitally extracted from a frame of the original video. As one particular example, Sitrick discloses how a face in an original video may be analyzed, after which a mask may be produced:

> "The analysis determines if a selected reference object appears in the visual picture image 710. In the example as depicted in FIG. 7, the visual picture image 710 includes *a reference object face 711 in the depicted example image frame*. In the example of FIG. 7, the face 711 is detected to be present. The tracking subsystem 700 may compute the location of the face 711 within the frame and output that information as position information 720." (*Id*. at ¶49.)

> "The tracking subsystem 700 *may compute [a]mask 750 which represents the region of the reference object within the visual picture image 710, in this example the face 711*. The mask may be output as a video output key signal. The mask may be output as a image alpha channel. The mask may be output in an encoded form. In a preferred embodiment, the mask is opaque in the region of the reference object and clear elsewhere. In another embodiment, the mask is clear in the region of the reference object and opaque elsewhere." (*Id*. at ¶54.)

FIG. 7 is reproduced below.




To form the mask 750, the "image is analyzed by the tracking subsystem 700 using general information known by the tracking subsystem 700 about the visual picture image 710." (*Id*. at ¶48.) A POSITA would know that analyzing the image 711

involves copying (extracting) pixel information so that the subsystem may determine the relative location of references objects, such as face 711, within the image 710. (Ex. 1003 at ¶93.) Since the mask is formed based on this extracted pixel information, Sitrick discloses the limitation of "wherein at least one pixel in a frame of said original video data stream is digitally extracted to form a first image."

Additionally, Sitrick discloses digitally extracting at least one pixel of the reference object to form a first image in relation to the use of image recognition. Specifically, Sitrick discloses various image recognition techniques used to analyze each frame of the original video data stream to detect "the reference object" (the first image) in a given frame. (Ex. 1007 at ¶71.) Sitrick discloses, with regard to Figs. 7 and 8, that "the face" 711/811 is the reference object in the image frame from the first audiovisual presentation. (*Id.* at ¶¶49, 57.) One known image recognition techniques disclosed in Sitrick is image matching.

> "Known forms of image recognition include image matching, where <u>*an image provided to the invention*</u> is compared and correlated against selected portions of the visual picture. The image is considered detected within the visual picture when the comparison and correlation exceed a threshold value." (*Id.* at ¶72 (emphasis added).)

Sitrick further discloses that the information about the reference object provided to the system may include the reference object's position within the visual image, a rotational orientation, color information, size information, geometric information such as a wire-frame mesh, mask information, and other information. (*Id.* at ¶82.)

Sitrick also discloses how a face (a reference object) in an original video may be analyzed:

> "*The analysis determines if a selected reference object appears in the visual picture image 710*. In the example as depicted in FIG. 7, *the visual picture image 710 includes a reference object face 711* in the depicted example image frame. In the example of FIG. 7, the face 711 is detected to be present." (*Id*. at ¶49 (emphasis added).)

A POSITA would understand that, in order to carry out the process (*e.g.*, image recognition) disclosed by Sitrick, the pixel information (*e.g.*, color information) relating to the selected reference object is extracted to form an image (the first image) wherein the disclosed system, using an image matching technique, may detect the first image in each frame of the original video data stream. (Ex. 1003 at ¶¶93-94.) In other words, the pixel information relating to the selected reference object is necessarily extracted from the original video data stream to form the first image so that the disclosed system, using an image matching

technique, may detect the first image in each frame of the original video data stream. Therefore, a POSITA would understand that Sitrick discloses forming the first image at least (1) when the mask is produced, or (2) when the image of the reference object is created to be used by the tracking subsystem. (*Id.*)

Sitrick further discloses that the first image (a mask or reference object) is replaced by second image thereon:

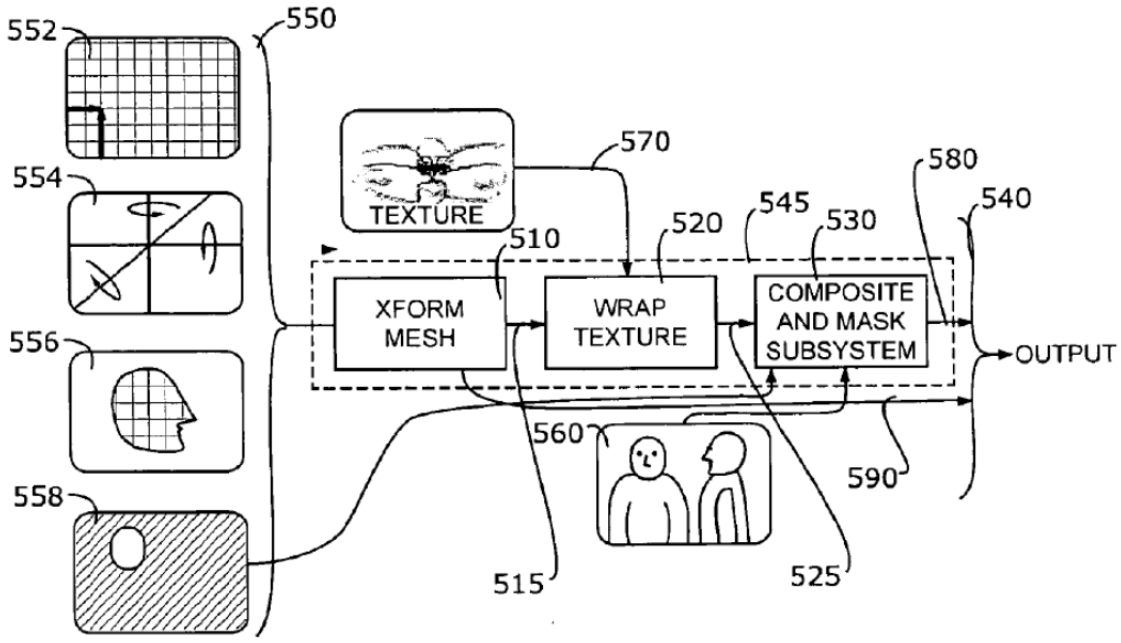
> "The invention then *replaces a portion of the first audiovisual presentation* with a portion of the associated replacement object image. The portion of the first audiovisual presentation selected is determined *by the associated reference object*. ((Ex. 1007 at ¶87 (emphasis added).)

A mask is one type of reference object. (*Id*. at ¶54.) Sitrick continues:

> It is not necessary to remove the selected portion of the first audiovisual presentation. *In a preferred embodiment, the portion of the associated replacement object image is overlaid on the reference object in the first audiovisual presentation*. The overlayment will obscure or replace a portion of the first audiovisual presentation, and is similar in nature to a video post-production effect commonly known as keying. (*Id*. at ¶87 (emphasis added).)

To obtain an image of the user's face, subsystem 100 processes the user's image data, such as in the manner illustrated in FIG. 5, reproduced below for convenience.

Fig. 5




Relative to FIG. 5, Sitrick discloses:

> "[0019] FIG. 5 is a system block diagram of a user image video processing and image integration subsystem of the present invention." (*Id*. at ¶19.)
>
> "In FIG. 5, the external source of user image content is shown representative of user image data comprising a texture map. The operation of the system shown in FIG. 5 is to use the position, rotation and orientation, and mesh geometry information present in the external program content to transform the mesh geometry information in the subsystem 510, producing a transformed mesh output on buses

> 515 and 590. . . *The use of the transform mesh subsystem coupled with the wrapping texture subsystem allows the subsystem 500 to recreate the appearance of the user from virtually any orientation or position by mapping the texture map onto the transformed mesh geometry. The compositing* ***and masking operation replaces*** *a selected portion of the program video 560 with the rendered image 525*. (*Id*. at ¶40 (emphasis added).)

Hence, Sitrick discloses "said first image then replaced by a second image resulting from a digital extraction of at least one pixel in a frame of a user input video data stream".

Thus, Sitrick discloses the limitation of [1-PREAMBLE-ii].

**[1a] an image capture device capturing the user input video data stream;**

Sitrick discloses that the user input video data stream (identified above) can be captured from a number of devices including a "video camera" or "digital camera." (*Id*. at ¶¶12, 139.) Thus, Sitrick discloses limitation [1a].

**[1b] an image display device displaying the original video stream;**

Sitrick discloses an image display device displaying the original video stream, such as the display device used to show the program video 120 illustrated in Fig. 1. *See also*, Figs. 2-6, showing similar program videos being displayed.

**[1c-i] a data entry device, operably coupled with the image capture device and the image display device;**

Sitrick discloses data entry devices meeting the requirements of limitation [1c-i]. Specifically, Sitrick discloses that the video compositing is preferably implemented via a general purposes computer:

> "FIG. 13 is a detailed block diagram of a preferred embodiment of the system of the present invention *implemented with a general purpose computer performing the compositing*." (*Id.* at ¶29.)

(*See also* ¶¶29, 41-43, 46, 69-70, 79-80, 95, 108-109, 115, 118, 121-122.) A general purpose computer would necessarily have a "data entry device," such as a keyboard. (Ex. 1003 at ¶¶98, 105.)

Sitrick's general purpose computer may be used to process both the user's image data and the program data. For instance, Fig. 1 includes an "integration subsystem" 100, which is used to create the output video 190 based on the program data 120 and the user image data 135. (Ex. 1007 at ¶31.) Further, the description of Fig. 13 states that the general purpose computer may be used to implement the compositing. (*Id.* at ¶29.) Thus, a POSITA would know that this integration subsystem could be implemented in Sitrick's general purpose computer. (Ex. 1003 at ¶99.) Therefore, Sitrick discloses, teaches or suggest the general purpose computer is the integration subsystem. (*Id.*)

As shown in Fig. 1, integration subsystem 100 is operably coupled to the external source of user image content having the user's image data. (Ex. 1007 at ¶31.) A POSITA would know that this image data would be provided by an "image capture device", such as by a "video camera." (Ex. 1003 at ¶100 (citing Ex. 1007 at ¶12).) Since the image capture device is operably connected to the general purpose computer, a POSITA would also know that the computer's data entry device (*e.g.*, a keyboard), could also be operably connected to the image capture device. Indeed, Sitrick discloses that "a user *selected* image is selectively integrated into a predefined audiovisual presentation…" (Ex. 1007 at ¶11), indicating that it is intended for the general purpose computer (and its applicable data entry devices) to interact with the image capture device so as to facilitate appropriate capture of the user's image(s). (Ex. 1003 at ¶100.) Thus, Sitrick discloses, teaches or suggests, a data entry device, operably coupled with the image capture device.

Sitrick also discloses that the general purpose computer is operably connected to the image display device, such as the display used to show the program video 120 of Fig. 1. Fig. 13 specifically provides for a video input signal 1315, which is used by the general purpose computer of Fig. 13 to complete the compositing. The video input signal 1315 is representative of the first audiovisual presentation (i.e., the program content 110 and its associated program video 120).

(Ex. 1003 at ¶101 citing (Ex. 1007 at ¶121).) Since the general purpose computer receiving the video input signal 1315 is coupled to the program content 110, a POSITA would also know that the computer's data entry device (e.g., a keyboard) could also be operably connected to the image display device. Indeed, Sitrick discloses that "[a]n analysis system analyzes the signals associated *with the selected portion of the predefined audiovisual presentation* …" (Ex. 1007 at ¶13), indicating to a POSITA that the user is using the data entry device to interact with the image display device via the program content 110 feed to the general purpose computer (integration subsystem). Thus, Sitrick also discloses, teaches or suggests "a data entry device, operably coupled with . . . the image display device."

Thus, Sitrick discloses, teaches or suggests the limitations of [1c-i].

**[1c-ii] . . . operated by a user to select the at least one pixel in the frame of the user input video data stream to use as the second image, and further operated by the user to select the at least one pixel to use as the first image;**

Building on the discussion of limitation [1c-i], above, Sitrick discloses that a user operates Sitrick's computer to "select the at least one pixel," for both the first and second images, as claimed:

> ". . . *a user selected image* is selectively integrated into a predefined audiovisual presentation in place of *a tracked portion* of the predefined audiovisual presentation. *A user can create a video or other image utilizing any one of a plurality of input device means*. . . ." (*Id.* at ¶11.)

Moreover, a user operating the Sitrick system would necessarily have to "select" at least one pixel in the original video data stream and the user input video data stream in order for the system to analyze "the selected portion of the predefined audiovisual presentation" (the first image) and "the user selected image" (the second image). (Ex. 1003 at ¶103 (citing Ex. 1007 at ¶13).)

Thus, Sitrick discloses limitation [1c-ii].

**[1d] wherein said data entry device is selected from a group of devices consisting of: a keyboard, a display, a wireless communication capability device, and an external memory device;**

Sitrick discloses this limitation. A general purpose computer necessarily includes a data entry device, such as a keyboard. (Ex. 1003 at ¶105.) Thus, Sitrick discloses limitation [1d].

**[1e] a digital processing unit operably coupled with the data entry device, said digital processing unit performing:**

Sitrick's general purpose computer includes normal computer components to carry out the functions described in Sitrick, including a CPU (a digital processing unit). (Ex. 1007 at ¶115.) The computer of Sitrick performs the stops required by limitations [1e-i] through [1e-vii], below. Thus, Sitrick discloses limitation [1e].

**[1e-i] identifying the selected at least one pixel in the frame of the user input video data stream;**

As explained above relative to the [1-PREAMBLE-ii] and [1c] limitations, Sitrick discloses selecting a user's face (second image) from the user's image data

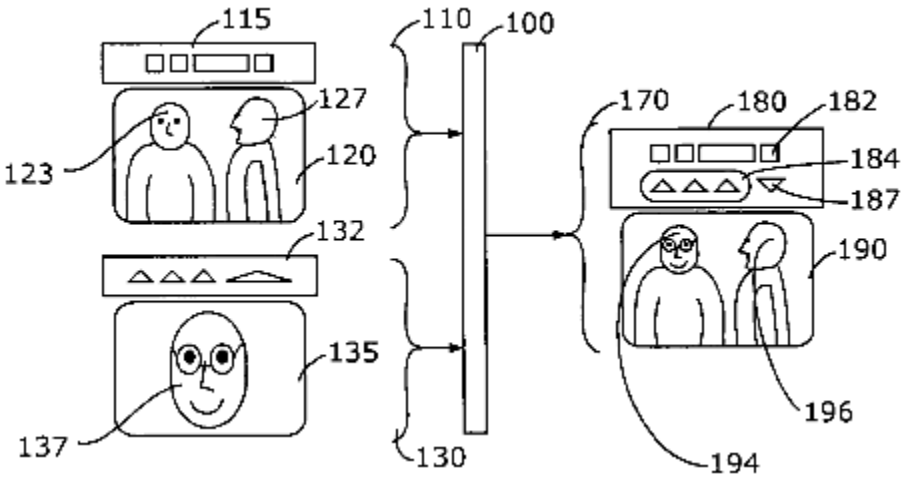
for overlaying on the mask/reference object (first image) of the program video. (*Id.* at ¶¶11, 31, 40, 87, Figs. 1 and 5.) To complete the overlaying, pixel(s) of the user image data must necessarily be identified and selected. (Ex. 1003 at ¶108.) Indeed, Sitrick discloses identifying various information (*e.g.*, "the user object geometric information and a pixel texture") relating to the selected at least one pixel in the frame of the user input video data stream. (Ex. 1007 at ¶104.)

Thus, Sitrick discloses limitation [1e-i].

**[1e-ii] extracting the identified at least one pixel as the second image;**

As shown in limitations [1c-ii] and [1e-i] above, Sitrick discloses that a replacement image (*e.g.*, a user's face) may be identified. Sitrick further discloses an extracted user selected image as a facial image 137:

Fig. 1




(*Id.* at Fig. 1.)

> "In the external source of user image content 130 is further comprised of other user data 132 and user image data 135, the user image data 135 is further comprised of *a user specified image 137*. In the figure, *137 appears as a single image of a face."* (*Id.* at ¶31.)

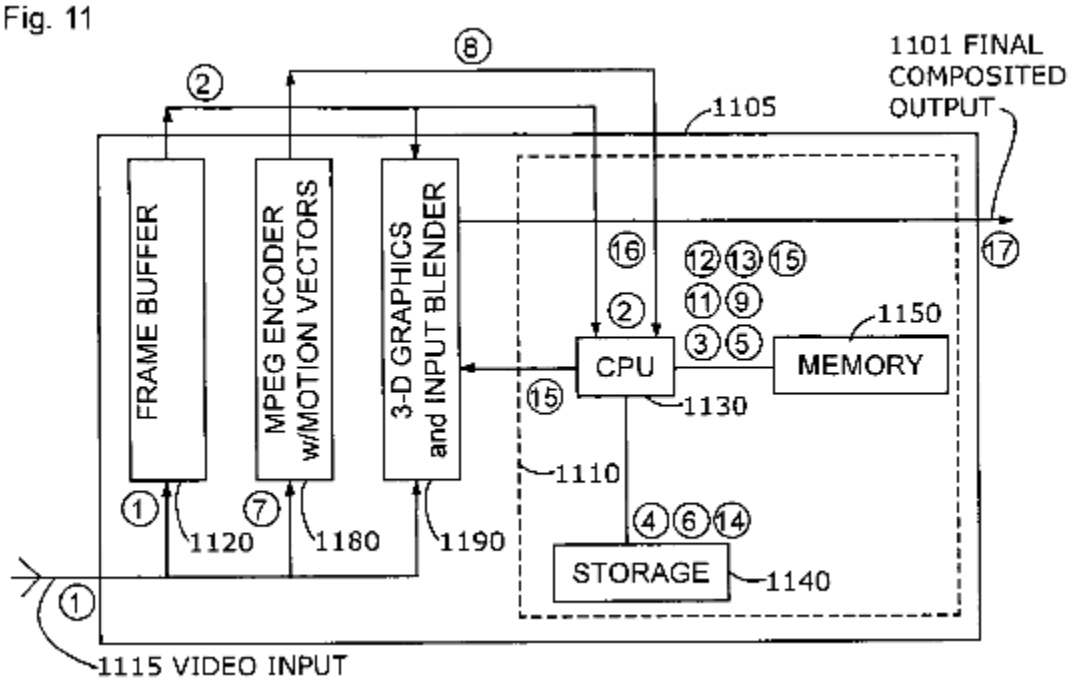
Sitrick further discloses that pixel information relating to the identified at least one pixel may be extracted from the user object. (*Id.* at ¶101.)

Accordingly, Sitrick discloses limitation [1e-ii].

**[1e-iii] storing the second image in a memory device operably coupled with the interactive media apparatus;**

As shown in limitation [1c-i], Sitrick discloses a general computing device having a memory and storage. (*Id.* at ¶115.) Sitrick further discloses storing the user's images in memory:

> "The database comprises a plurality of user replacement object images 931, 932, 933, 934, 935, and 936. The database further comprises a plurality of user geometric information data structures 941, 942, 943, 944, 945, and 946." (*Id.* at ¶111.)
>
> "The data for the user replacement object image may reside in either or both of the storage subsystem 1140 or the memory subsystem 1150." (*Id.* at ¶116.)




(*Id.* at Fig. 11.)

Thus, Sitrick discloses limitation [1e-iii].

**[1e-iv] receiving a selection of the first image from the original video data stream;**

As described above relative to claim limitation [1c-ii], Sitrick describes producing a first image (a mask or reference object) that would be replaced with a user selected image (the second image). Thus, a user operating the Sitrick system "selects" the first image so that the system analyzes "the selected portion of the predefined audiovisual presentation." (Ex. 1003 at ¶111 (citing Ex. 1007 at ¶13).) Indeed, Sitrick discloses that, in order to carry out the disclosed process (*e.g.*, tracking), the reference object (the first image) has to be identified first by the

system. (Ex. 1007 at ¶84.) Accordingly, the Sitrick system, which may be implemented on a general purpose computer, necessarily receives the selection of the first image in order to carry out the disclosed replacement process. (Ex. 1003 at ¶111 (citing Ex. 1007 at ¶115).)

Thus, Sitrick discloses limitation [1e-iv].

**[1e-v] extracting the first image;**

As explained above in the [1-PREAMBLE-ii] limitation, Sitrick extracts a first image, such as a mask or reference object image. (*Id*. at, *inter alia*, Fig. 7 and ¶¶48-49, 54 (shows extraction of the mask image); Figs. 7-8, ¶¶49, 57, 71-72, 82 (shows extraction of a reference object image).) (Ex. 1003 at ¶¶112-114 (explaining Sitrick's extraction of the mask and reference object images).).

Thus, Sitrick discloses limitation [1e-v].

**[1e-vi] spatially matching an area of the second image to an area of the first image in the original video data stream, wherein spatially matching the areas results in equal spatial lengths and widths between said two spatially matched areas; and**

Sitrick discloses limitation 1e-vi. In order to overlay the mask/reference object with the second image, the images must necessarily be spatially matched in the X-Y dimensions (length-width). Sitrick discloses several methods of matching an area of the second image to an area of the first image—*e.g.*, mapping, stretching, rotating, scaling, zooming, curling, shearing, distorting, and morphing of the size

of a replacement image (second image) to obtain the best results. (Ex. 1007 at ¶¶94-96, 100.)

Accordingly, Sitrick discloses limitation 1e-vi.

**[1e-vii] performing a substitution of the spatially matched first image with the spatially matched second image to generate the displayable edited video data stream from the original video data stream**

As discussed in limitation [1e-v] and [1e-vi], above, the second image is substantially matched to a portion of the first audiovisual presentation (the mask/reference object or first image) and the second image overlays (is substituted for) the first image. (*Id.* at ¶¶31, 87, 95-96, 100.) Also, as discussed in Overview of Sitrick, the system creates an edited video data stream (output video 190) by performing the replacement. (*Id.* at Fig. 1.) Accordingly, Sitrick discloses limitation [1e-vii].

Thus, Sitrick discloses, teaches, or suggests all limitations of claim 1, and therefore renders claim 1 obvious.

**[Claim 2] The interactive media apparatus of claim 1 wherein the digital processing unit is further capable of performing: computing motion vectors associated with the first image; and applying the motion vectors to the second image extracted from the user input video data stream, wherein the generated displayable edited video data stream resulting from the substitution maintains an overall motion of the original video data stream:**

Sitrick discloses that its computer computes the location of the reference object in each frame of a video:

> "The analysis *determines if a selected reference object appears in each image in the time-ordered sequence 810*. . . . The tracking subsystem 800 may *compute the location of the face 811 within the frame 815* and output that information as position information 830." (*Id.* at ¶57.)

Sitrick further discloses that the video may be encoded using the MPEG standard, where the encoded video contains motion vector information used by a "correlation means." [8] (*Id*. at ¶65,67.) Accordingly, a POSITA understands Sitrick as disclosing that its computer computes the motion vectors in a video encoded in the MPEG standard to estimate the actual position of the reference object in each frame of the video. (Ex. 1003 at ¶119.)

Sitrick further discloses that the information obtained through the motion vectors information associated with the first image is applied to the second image by geometrical transformation to match-up the first and second images. (Ex. 1007 at ¶¶100, 104.) A POSITA knows that, because the motion vectors associated with the first image are applied to the second image, the motion vectors associated with

[8] Petitioner notes that MPEG stands for "Moving Picture Experts Group." Exhibit 1015.

the first image will be maintained in the edited video data stream. (Ex. 1003 at ¶120.)

Thus, Sitrick discloses all limitations of claim 2, and renders obvious claim 2.

**[Claim 8] The interactive media apparatus of claim 1, wherein the substitution performed by the digital processing device replaces at least a face of a first person from the original video data stream by at least a face of a second person from the user input video data stream.**

Sitrick discloses this limitation as the first person's facial image 123 is replaced with a second person's facial image 137. (Ex. 1007 at ¶31, Fig. 1.) Accordingly, Sitrick renders obvious claim 8.

**[Claim 11]**

Claim 11 includes the same limitation as claim 1, but is written as a method claim instead of an apparatus claim. *See* VII.A.2.[Claim 11] above for the table comparing claim limitations of claims 1, 2, and 11. For the same reasons Sitrick makes obvious claims 1-2, Sitrick also renders obvious claim 11.

### D. Ground 4: Claims 3 and 4 are rendered obvious by Sitrick in view of Levoy

#### 1. Scope and Content of the Prior Art

For Ground 2, the prior art consists of Senftner and Levoy. Sitrick is described in detail above in Ground 3, and is pre-AIA 102(b) prior art to the '591 patent. Levoy is described in detail above in Ground 2, and is pre-AIA §§102(a),(e)

prior art to the ’591 patent. As discussed in Ground 2 (VII.B.1), Levoy describes systems and methods for selecting one or more portions of an image via a touch screen device, and then incorporating those selected portions into another image to creating a new composite image, such as one a mobile phone or other touch screen device.

**2. Motivation to Combine**

Both Sitrick and Levoy are directed to the same fundamental process of, while using a computing device, substituting an image data with another image data, creating a new composite data. While Sitrick broadly discloses that a general purpose computer may be used in preferred embodiments, Sitrick does not disclose any device that inherently has a touch screen. (*Id.* at ¶115.) However, Sitrick also does not limit the disclosed invention to a specific set of user input devices; instead, it describes that a user may utilize “any one of a plurality of input device means.” (*Id.* at ¶11.)

As described above in VII.B.2, touch screen devices and their benefits were well-known at the time of the ’591 patent was filed. Thus, a POSITA would find it obvious to use conventional touch screen technology with the general purpose computer recited by Sitrick. (Ex. 1003 at ¶142.)

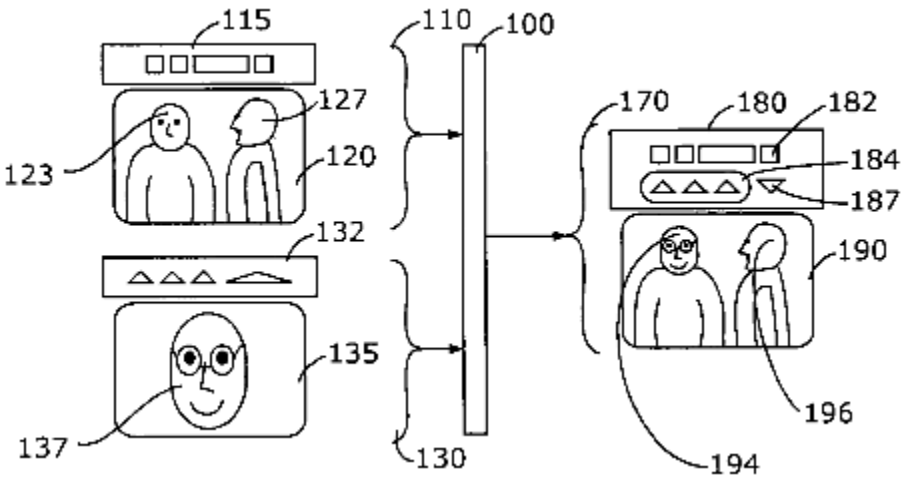
### 3. Claims

**[Claim 3] The interactive media apparatus of claim 1 wherein the digital processing unit is further capable of extracting the at least one pixel from the user entering data in the data entry display device.**

As discussed in detail in VII.B.3.[Claim 3] above, Petitioner notes that Claim 3 is invalid on multiple §112 grounds. Nonetheless, for the purposes of this IPR only, Petitioner assumes that "data display entry device" of claim 3 means the "data entry device" of claim 1 is a display and that the "at least one pixel" of claim 3 is construed to refer to the "user input data stream" of claim 1.

Assuming claim 3 can be construed, Sitrick in view of Levoy renders claim 3 obvious. As shown above relative to claim 1, Sitrick discloses that a replacement image (*e.g.*, a user's face) may be identified and discloses an extracted user selected image as a facial image 137. (Ex. 1007 at Fig. 1 and ¶31.)

Fig. 1




(*Id.* at Fig. 1.)

Sitrick further discloses that pixel information relating to the identified at least one pixel may be extracted from the user object. (*Id.* at ¶101.) Claim 3 further limits claim 1 by requiring that the user enter the data "in the data display entry device." As noted above, Petitioner assumes "the data display entry device" means the "data entry device" is a display, such as a touch screen.

Sitrick discloses a general purpose computer can be used to select the original images and new images for replacement. (*Id.* at ¶¶11, 13, Fig. 1.) Sitrick's computer necessarily includes a data entry device, but Sitrick does not specifically disclose a touch screen data entry device. (Ex. 1003 at ¶105.)

However, described in the Motivation to Combine section (VII.D.II), above and as evidenced by Levoy, touch screen devices were generally well-known at the time of the '591 patent, and were known to provide benefits over non-touch screen devices. Thus, instead of using a conventional non-touch screen computer, a POSITA would find it obvious to use conventional touch screen technology with the general purpose computing device of Sitrick, which would allow a user to copy (extract) pixels from the user entering data in the touch screen (data entry display device). (*Id*. at ¶145.)

Thus, Sitrick in view of Levoy teaches or suggests the limitations of claim 3, rendering claim 3 obvious.

**[Claim 4] The interactive media apparatus of claim 3 wherein the digital processing unit is further capable of extracting the at least one pixel from the user pointing to a spatial location in a displayed video frame**

Like claim 3, claim 4 is invalid under §112 for lack of written description and for being indefinite. Nonetheless, for purposes of this IPR only, Petitioner shows that claim 4 is invalid when the "at least one pixel" of claim 4 is construed to refer to the "user input data stream" of claim 1.

Sitrick in view of Levoy teaches or suggests the limitations of claim 4. As described above relative to claim 3, Sitrick does not specifically disclose using a touch screen device as the computing device. As also described above relative to claim 3, it would be obvious in view of Levoy to use a touch screen computing device as the computing device of Sitrick. When using touch screen technology to edit pictures, a user would naturally point "to a spatial location in a displayed video frame", as evidenced by Levoy. (Ex. 1008 at ¶47.)

Thus, Sitrick in view of Levoy renders claim 4 obvious.

## VIII. CONCLUSION

For the foregoing reasons, Petitioner respectfully requests that the *inter partes* review of the '591 patent be instituted as the Petition establishes a reasonable likelihood of prevailing with respect to the Challenged Claims. Petitioner further respectfully requests that Claims 1-4, 8, and 11 be cancelled as unpatentable.

Respectfully submitted,
GREENBERG TRAURIG, LLP

Date: March 29, 2017

By: */s/ Heath J. Briggs*

Heath J. Briggs
Registration No. 54,919
1200 17th Street, Suite 2400
Denver, CO 80202
Phone: (303) 572-6500
Fax: (303) 572-6540
*Counsel for Petitioner*

## CERTIFICATE OF SERVICE

The undersigned certifies that a true and correct copy of the Petition together with all exhibits identified in the above Table of Exhibits and Petitioner's Power of Attorney, have been served on the Patent Owner via Priority Mail Express or by means at least as fast and reliable as Priority Mail Express on the below date, at the following addresses:

Michael J. Buchenhorner
8540 S.W. 83 Street
Miami, FL 33143

Respectfully submitted,
GREENBERG TRAURIG, LLP

Date: March 29, 2017

By: */s/ Heath J. Briggs*
Heath J. Briggs
Registration No. 54,919
1200 17th Street, Suite 2400
Denver, CO 80202
Phone: (303) 572-6500
Fax: (303) 572-6540
*Counsel for Petitioner*