(51) **International Patent Classification:**
G06K 9/36 (2006.01)

(21) **International Application Number:**
PCT/US2010/000009

(22) **International Filing Date:**
5 January 2010 (05.01.2010)

(25) **Filing Language:** English

(26) **Publication Language:** English

(30) **Priority Data:**
61/207,261   10 February 2009 (10.02.2009)   US

(71) **Applicant** *(for all designated States except US)*: **THOMSON LICENSING** [FR/FR]; 1 rue Jeanne d'Arc, F-92443 Issy-les moulineaux (FR).

(72) **Inventors; and**

(75) **Inventors/Applicants** *(for US only)*: **BHAGAVATHY, Sitaram** [IN/US]; 5910 Hunters Glen Drive, Plainsboro, New Jersey 08536 (US). **LLACH, Joan** [ES/US]; 101 Castleton Road, Princeton, New Jersey 08540 (US).

(74) **Agents:** **SHEDD, Robert, D.** et al.; Thomson Licensing LLC, Two Independence Way, Suite 200, Princeton, New Jersey 08540 (US).

(81) **Designated States** *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) **Designated States** *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

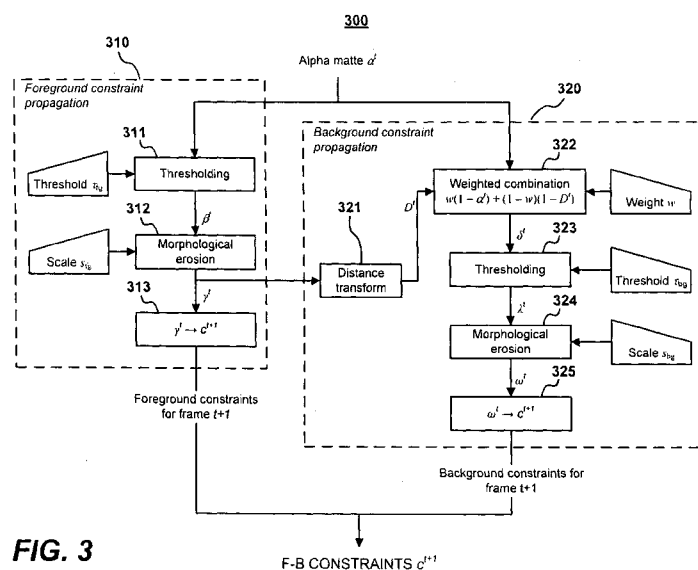(54) **Title:** VIDEO MATTING BASED ON FOREGROUND-BACKGROUND CONSTRAINT PROPAGATION

FIG. 3

(57) **Abstract:** A method for propagating user-provided foreground-background constraint information for a first video frame to subsequent frames allows extraction of moving foreground objects with minimal user interaction. Video matting is performed wherein constraints derived from user input with respect to a first frame are propagated to subsequent frames using the estimated alpha matte of each frame. The matte of a frame is processed in order to arrive at a rough foreground-background segmentation which is then used for estimating the matte of the next frame. At each frame, the propagated constraints are used by an image matting method for estimating the corresponding matte which is in turn used for propagating the constraints to the next frame, and so on.

VIDEO MATTING BASED ON FOREGROUND-BACKGROUND
CONSTRAINT PROPAGATION

Related Patent Applications

[0001]   This application claims the benefit under 35 U.S.C. § 119(e) of United States Provisional Application No. 61/207,261, filed February 10, 2009, the entire contents of which are hereby incorporated by reference for all purposes into this application.

Field of Invention

[0002]   The present invention generally relates to digital image processing, and more particularly to matting methods and apparatus for video images.

Background

[0003]   Image matting is the process of extracting an object from an image with some human guidance. Image matting may be an interactive process which relies on limited user input, usually in the form of a few scribbles, to mark foreground and background regions.   Henceforth, "foreground" refers to the object to be extracted, whereas "background" refers to everything else in the image.

[0004]   Video matting is an extension of image matting wherein the goal is to extract a moving object from a video sequence. Video matting can also be used in video processing devices (including video encoders). For instance, automatic matte extraction can be used to identify a particular region in a video scene (e.g. sky area), and then apply a given processing only to that region (e.g. de-banding or false contour removal). Matte extraction can also be used to guide object detection and object tracking algorithms. For instance, a matte extraction technique could be used to detect the grass area in a soccer video (i.e. the playfield) which could then use to constrain the search range in a ball tracking algorithm.

[0005]   In moviemaking and television, mattes have been used to composite foreground (e.g. actors) and background (e.g. landscape) images into a final image. The chroma keying (blue screen) technique is a widely used method for matting actors into a novel background. Many of the traditional techniques rely on a controlled environment during the image capture process. With digital images, however, it becomes possible to

directly manipulate pixels, and thus matte out foreground objects from existing images with some human guidance. Digital image matting is used in many image and video editing applications for extracting foreground objects and possibly for compositing several objects into a final image.

5   [0006]   As mentioned, image matting is usually an interactive process in which the user provides some input such as marking the foreground and possibly the background regions. The simpler the markings are, the more user-friendly the process is. Among the easier-to-use interfaces are those in which the user places a few scribbles with a digital brush marking the foreground and background regions (see FIG. 2A). An image matting
10   process then determines the boundary of the foreground object using the image information along with the user input.

[0007]   In several image matting methods, the user provides a rough, usually hand-drawn, segmentation called a trimap, wherein each pixel is labeled as a foreground, background, or unknown pixel. (See U.S. Patent No. 6,135,345 to Berman et al.,
15   "Comprehensive method for removing from an image the background surrounding a selected object"; and Y. Y. Chuang et al., "A Bayesian approach to digital matting," Proc. IEEE Conf. on Computer Vision and Pattern Recognition, 2001.) Other methods allow a more user-friendly scribble-based interaction in which the user places a few scribbles with a digital brush marking the foreground and background regions. (See J.
20   Wang et al., "An iterative optimization approach for unified image segmentation and matting," Proc. IEEE Conf. on Computer Vision and Pattern Recognition, 2005; and A. Levin et al., "A closed-form solution to natural image matting," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 30, no. 2, pp. 228-242, Feb 2008.)

[0008]   In all the above methods, the user input is provided for matting out the
25   foreground from a single image. Video matting is a harder problem as it may involve a moving foreground object. In this case, the user input for one frame may not be accurate for subsequent frames. Moreover, it is labor-intensive to require the user to provide input for each frame in the video.

[0009]   In the video matting method proposed by Chuang et al., a trimap is provided
30   for each of several keyframes in the video, and the trimaps are interpolated to other frames using forward and backward optical flow. (Y. Y. Chuang et al., "Video matting of

complex scenes," ACM Transactions on Graphics, vol. 21, no. 3, pp. 243-248, 2002.)
Optical flow-based interpolation, however, is time-consuming, noise sensitive, and
unreliable, even for moderate motion levels. Furthermore, optical flow-based
interpolation of user-provided scribbles results in the scribbles breaking up over time.

5   Apostoloff et al. describe a method in which trimaps are implicitly propagated from
frame to frame by imposing spatiotemporal consistency at edges. (N. E. Apostoloff et al.,
"Bayesian video matting using learnt image priors," Proc. IEEE Conf. on Computer
Vision and Pattern Recognition, 2004.) The complexity of this method, however, can be
substantial due to the enforcement of spatiotemporal edge consistency between the
10   original image and the alpha mattes

Summary

[0010] In an exemplary embodiment in accordance with the principles of the
invention, a method is described for propagating user-provided foreground-background
15   constraint information for a first video frame to subsequent frames, thereby allowing
extraction of moving foreground objects in a video stream with minimal user interaction.
Video matting is performed wherein the user input (e.g. scribbles) with respect to a first
frame is propagated to subsequent frames using the estimated matte of each frame. The
matte of a frame is processed in order to arrive at a rough foreground-background
20   segmentation which is then used for estimating the matte of the next frame. At each
frame, the propagated input is used by an image matting method for estimating the
corresponding matte which is in turn used for propagating the input to the next frame, and
so on.

[0011] In view of the above, and as will be apparent from the detailed description,
25   other embodiments and features are also possible and fall within the principles of the
invention.

Brief Description of the Figures

[0012] Some embodiments of apparatus and/or methods in accordance with
30   embodiments of the present invention are now described, by way of example only, and
with reference to the accompanying figures in which:

[0013]   FIG. 1 is a block diagram of an exemplary frame-wise matting process or apparatus;

[0014]   FIGs. 2A-2D illustrate an interactive image matting process, in which FIG. 2A shows an image or video frame with white scribbles marking foreground and black scribbles marking background; FIG. 2B shows an extracted matte or foreground opacity for FIG. 2A; FIG. 2C shows an extracted foreground image; and FIG. 2D shows the foreground object composited with a novel background image;

[0015]   FIG. 3 is a flow diagram of a foreground-background constraint propagation method;

[0016]   FIGs. 4A through 4F show illustrative images pertaining to the method of FIG. 3; and

[0017]   FIG. 5 is a block diagram of an exemplary system embodiment of the present invention.

Description of Embodiments

[0018]   Other than the inventive concept, the elements shown in the figures are well known and will not be described in detail. For example, other than the inventive concept, familiarity with digital image processing is assumed and not described herein. It should also be noted that embodiments of the invention may be implemented using various combinations of hardware and software. Finally, like-numbers in the figures represent similar elements.

[0019]   In general, image matting methods process an input image I which is assumed to be a composite of a foreground image F and a background image B. The color of the ith pixel is assumed to be a linear combination of the corresponding foreground F_i and background B_i colors or intensities:

$$I_i = \alpha_i F_i + (1 - \alpha_i) B_i,\tag{1}$$

where α_i is the foreground opacity of the ith pixel and $0 \le \alpha \le 1$. Not all of the quantities on the right hand side of Eq. 1 are known. Thus, for a three-channel color image, there are three equations and seven unknowns for each pixel of the image. Because this is an under-constrained problem, some user input is required to extract a good matte. Typical user input for this purpose may include, for example, placing a few scribbles with a

digital brush marking the foreground and background regions, as illustrated in FIG. 2A. In FIG. 2A, the white scribbles denote foreground regions and the black scribbles denote background regions. Given the user input, the foreground image F, the background image B, and the alpha matte (or just "matte") α are estimated.

5   [0020]   FIG. 1 is a block diagram of an exemplary frame-wise matting process or apparatus in which the foreground is extracted on a frame-by-frame basis using image matting block 110. As shown in FIG. 1, for a frame at time t, image matting block 110 generates an associated alpha matte α^t based on foreground-background (F-B) constraints c^t that are based on input by a user for frame t or propagated from user input

10   for a previous frame via constraint propagation block 120. In frame t, the constraints c^t are determined by the user-provided input designating foreground and background regions, where all pixels marked as foreground have $\alpha = 1$ and all pixels marked as background have $\alpha = 0$.

[0021]   Referring to FIG. 1, c^t denotes the F-B constraints of frame t. For the ith pixel,

15
$$c_i^t = \begin{cases} 1 & \text{if pixel } i \text{ is marked as foreground} \\ 0 & \text{if pixel } i \text{ is marked as background} \\ \text{undefined} & \text{otherwise} \end{cases} \qquad (2)$$

[0022]   The image matting method 110 assumes the constraints $\alpha_i^t = 1$ for pixels with $c_i^t = 1$ and $\alpha_i^t = 0$ for pixels with $c_i^t = 0$, where α_i^t denotes the alpha value of the ith pixel in frame t. FIG. 2B shows the matte (α^t) obtained for frame t based on the user-provided scribbles (c^t) in FIG. 2A. FIG. 2C shows the extracted foreground object and FIG. 2D

20   shows an application of the matting in which the extracted foreground object is composited with a novel background image.

[0023]   In an exemplary embodiment, matting block 110 can be implemented in accordance with the matting technique described in A. Levin et al., "A closed-form solution to natural image matting," IEEE Trans. on Pattern Analysis and Machine

25   Intelligence, vol. 30, no. 2, pp. 228-242, Feb 2008. The matting technique of Levin et al. minimizes a quadratic cost function in the alpha matte under some constraints. Any suitable matting techniques may be used for this purpose.

[0024]   For each subsequent frame in a video stream, the foreground-background constraints can be derived by propagating the user-input constraints from a previous frame via constraint propagation block 120. An exemplary method for propagating the foreground-background (F-B) constraints from one frame to the next will now be described.

[0025]   FIG. 3 is a flow diagram of an exemplary method 300 of propagating F-B constraints from one frame to the next is outlined in FIG. 3. FIGs. 4A through 4F show illustrative images pertaining to the method 300. As described in greater detail below, the method 300 uses the alpha matte at frame t to estimate F-B constraints at time $t+1$. Instead of propagating the constraints (e.g., scribbles or trimap) directly, the method uses the matte at t to generate the constraints for frame $t+1$. In other words, using the matte (α) from frame t, the F-B constraints (c^{t+1}) for the frame $t+1$ are obtained.

[0026]   As shown in FIG. 3, the alpha matte α_i^t for the first of a sequence of frames, i.e., for the frame of time t, is provided, such as from the matting method of FIG. 1, as an input to the method 300. The method 300 comprises foreground constraint propagation procedure 310 and background constraint propagation procedure 320.

[0027]   The alpha value α_i^t for each pixel i lies in the range [0, 1], where $\alpha_i^t = 1$ means that pixel i entirely belongs to the foreground. The higher its alpha value, the greater the foreground component of the pixel. At step 311, a thresholding operation is performed in which the alpha values are compared to a threshold τ_{fg} to generate a binary field β^t such that:

$$\beta_i^t = \begin{cases} 1 & \text{if } \alpha_i^t > \tau_{fg} \\ 0 & \text{otherwise} \end{cases} . \tag{3}$$

The thresholding operation thus isolates pixels with a high foreground component. FIG. 4A illustrates the results of the thresholding step 311. FIG. 4A shows the binary field β^t obtained by thresholding the alpha matte α^t shown in FIG. 2B. In various exemplary embodiments, values of threshold τ_{fg} are in the range of 0.5 to 0.9.

[0028]   At step 312, morphological erosion is performed on the binary field β^t generated by the thresholding step:

$$\gamma' = F_E\left(\beta', E(s_{f_g})\right), \tag{4}$$

where $F_E(.)$ denotes the morphological erosion operator and $E(s)$ denotes a structuring element of scale s. In an exemplary embodiment, $E(5)$ is a disk with a radius of 5 pixels. The structuring element can be any suitable shape, including for example, a square, however, an isotropic shape such as a disk is preferable. The scale of the structuring element is preferably selected based on the desired separation between the foreground and background, and the size of the foreground and/or background, so as avoid excessive erosion of foreground and/or background pixels.

[0029]   Note that if thresholding step 311 yields a small foreground area that would be eliminated or reduced by morphological erosion step 312 to a foreground area smaller than a predetermined minimum size (such as a size too small to be perceived by a viewer), morphological erosion step 312 may be skipped.

[0030]   FIG. 4B illustrates the field γ' resulting from the morphological erosion of binary field β' in step 312. In FIG. 4B, the foreground constraints are indicated by the white pixels. By shrinking the regions marked as foreground, the erosion operation ensures that these regions lie within the foreground in frame $t+1$ even if the foreground (and/or background) has moved by a certain amount. The scale s_{f_g} of the structuring element $E(.)$ can be chosen based on the degree of motion of the foreground and/or background.

[0031]   The foreground constraints for the frame $t+1$ are then defined at step 313 as follows: $c_i^{t+1} = 1$ if $\gamma_i' = 1$. If $\gamma_i' = 0$, c_i^{t+1} is yet undefined. Note that c_i^{t+1} may be set as background at another point in the process, as described below, or remain undefined.

[0032]   In the background propagation procedure 320, the background constraints are determined based on the already-determined foreground constraints γ_i' (FIG. 4B). For each pixel, a lower alpha value (α_i') indicates a higher background component. By applying a threshold on $(1 - \alpha_i')$, the background pixels may be isolated. However, it is not desirable that the background-constrained pixels lie very far from the foreground, since this would increase the "fuzzy zone" between background-constrained and foreground-constrained pixels, possibly resulting in inaccurate foreground extraction. In

order to achieve this balance, a normalized distance transform is determined at step 321 in accordance with the following expression:

$$D_i' = \frac{\min_{j \in J} d(i,j)}{\max(H,W)},$$ (5)

where $d(i,j)$ is the spatial distance between pixels i and j, J is the set of all pixels j with γ_j'

5      $= 1$, i.e. $J = \{j \mid \gamma_j' = 1\}$, and H and W are the height and width of the frame.

[0033]   Using the distance transform determined in step 321, a background score of each pixel i is determined at step 322 as a weighted combination of the inverses or complements of the alpha matte and the normalized distance transform, as follows:

$$\delta_i' = w(1 - \alpha_i') + (1 - w)(1 - D_i').$$ (6)

10   In an exemplary embodiment, the weight w has a value of 0.8. In various exemplary embodiments, the weight w has a range of 0.5 to 0.9.

[0034]   The weighted combination of Eq. 6 yields higher background scores for pixels that have a low alpha value and are situated close to foreground pixels. FIG. 4C illustrates the background score field δ' determined in step 322. The brighter the pixel

15   appears in FIG. 4C, the higher its background score.

[0035]   The background score determined in step 322 is then subjected to a thresholding operation in step 323 in which the background score field δ' is compared to a threshold to generate a binary field λ' such that:

$$\lambda_i' = \begin{cases} 1 & \text{if } \delta_i' > \tau_{\text{bg}} \text{ and } \gamma_i' = 0 \\ 0 & \text{otherwise} \end{cases},$$ (7)

20   where τ_{bg} is a preset background score threshold. An exemplary range of values for threshold τ_{bg} is 0.5 to 0.9.

[0036]   The binary field λ' generated in step 323 is then morphologically eroded in step 324:

$$\omega' = F_E\left(\lambda', E(s_{\text{bg}})\right)$$ (8)

25   where $F_E(.)$ denotes the morphological erosion operator and $E(s)$ denotes a structuring element of scale s. A variety of structuring elements of various shapes and sizes can be

9

used, as discussed above.   FIG. 4D illustrates the field ω' resulting from the morphological erosion of binary field λ' in step 324.   The white pixels in FIG. 4D indicate the background-constrained pixels.   By shrinking the regions marked as background, the erosion operation ensures that these regions lie outside the foreground in frame $t+1$ even if the foreground (and/or background) has moved by a certain amount.

[0037]   Note that if thresholding step 323 yields a small background area that would be eliminated or reduced by morphological erosion step 324 to a background area smaller than a predetermined minimum size (such as a size too small to be perceived by a viewer), morphological erosion step 324 may be skipped.

[0038]   Finally, at step 325, the background constraints for the frame $t+1$ are determined as follows: $c_i^{t+1} = 0$ if $\omega_i' = 1$. Any c_i^{t+1} that has not already been set to 0 or 1 is left undefined; i.e., such a pixel has an unknown constraint.

[0039]   FIG. 4E illustrates the propagated foreground-background constraints for the frame $t+1$ where white represents foreground constraints ($c_i^{t+1} = 1$) and black represents background constraints ($c_i^{t+1} = 0$).   FIG. 4F illustrates the matte α^{t+1} extracted from the frame $t+1$ using the propagated constraints.   The matte α^{t+1} can be generated using the matting method 110 of FIG. 1. This matte is in turn used for deriving the constraints for frame $t+2$ and so on.

[0040]   The exemplary method avoids the complexity of motion estimation methods such as correlation-based template matching or optical flow and works reliably over a range of motion levels.

[0041]   In an exemplary embodiment of a method of propagating F-B constraints from one frame to the next, prior information such as the area of the foreground object and its color distribution in the current frame is used in deriving the F-B constraints for the next frame.   All or a subset of the parameters τ_{fg}, τ_{bg}, s_{fg}, s_{bg}, and w can be automatically adjusted based on the prior information in order to extract an accurate matte. This process can be carried out iteratively until the matte satisfies the constraints imposed by the prior information. In an exemplary embodiment, a brute force process includes trying out multiple values, preferably within predefined ranges, for each parameter and selecting the set of values that best satisfies the prior information. As an example, consider an embodiment in which the prior information includes the area of the

foreground. If the foreground constrained area, such as determined by the above-described procedure, is too large, the parameters τ_{fg}, and s_{fg} can be increased, as this will result in fewer foreground constrained pixels. These parameters can be adjusted until the prior constraints are satisfied.

5 **[0042]**   FIG. 5 is a block diagram of an exemplary system 500 in accordance with the principles of the invention. The system 500 can be used to generate alpha mattes, F-B constraints, and/or perform matting from a video stream. The system 500 comprises a frame grabber 510 and a digital video editor 520. Frame grabber 510 captures one or more frames of the video stream for processing by digital video editor 520 in accordance

10 with the principles of the invention. Digital video editor 520 comprises processor 521, memory 522 and I/O 523. In an exemplary embodiment, digital video editor 520 may be implemented as a general purpose computer executing software loaded in memory 522 for carrying out constraint propagation and/or matting as described above.

15 **[0043]**   In view of the above, the foregoing merely illustrates the principles of the invention and it will thus be appreciated that those skilled in the art will be able to devise numerous alternative arrangements which, although not explicitly described herein, embody the principles of the invention and are within its spirit and scope. For example, although illustrated in the context of separate functional elements, these functional

20 elements may be embodied in one, or more, integrated circuits (ICs). Similarly, although shown as separate elements, some or all of the elements may be implemented in a stored-program-controlled processor, e.g., a digital signal processor or a general purpose processor, which executes associated software, e.g., corresponding to one, or more, steps, which software may be embodied in any of a variety of suitable storage media. Further,

25 the principles of the invention are applicable to various types of wired and wireless communications systems, e.g., terrestrial broadcast, satellite, Wireless-Fidelity (Wi-Fi), cellular, etc. Indeed, the inventive concept is also applicable to stationary or mobile receivers. It is therefore to be understood that numerous modifications may be made to the illustrative embodiments and that other arrangements may be devised without

30 departing from the spirit and scope of the present invention.

<u>CLAIMS</u>

1.   A computer implemented method for determining foreground and background constraints for a video stream comprising:

5   receiving foreground and background constraints for a video frame;

          generating a matte for the video frame using the foreground and background constraints;

          generating foreground constraints for a subsequent video frame using the matte; and

10   generating background constraints for the subsequent video frame using the matte.

2.   The method of claim 1, wherein the foreground and background constraints for the video frame are received via user input.

15

3.   The method of claim 1, wherein generating foreground constraints for the subsequent video frame includes:

          performing a thresholding operation on the matte to generate a binary field.

20 4.   The method of claim 3, wherein generating foreground constraints for the subsequent video frame includes:

          performing a morphological erosion operation on the binary field.

5.   The method of claim 1, wherein generating background constraints for the

25 subsequent video frame includes:

          generating a background score field using the matte; and

          performing a thresholding operation on the background score field to generate a binary field.

6.   The method of claim 5, wherein generating background constraints for the subsequent video frame includes:

performing a morphological erosion operation on the binary field.

5

7.   The method of claim 5, wherein the background score field includes background scores for pixels of the video frame, the background score for a pixel being directly related to the pixel's proximity to a foreground pixel.

10   8.   The method of claim 5, wherein the background score field includes background scores for pixels of the video frame, the background score for a pixel being inversely related to a corresponding value for the pixel in the matte.

9.   The method of claim 5, wherein generating the background score field includes:

15   generating a distance transform from the foreground constraints; and

generating a weighted combination of an inverse of the matte and an inverse of the distance transform.

10.   The method of claim 9, wherein generating the weighted combination includes

20   weighting the inverse of the matte by a weight w and the inverse of the distance transform by a weight $1-w$, wherein $0.5 \leq w \leq 0.9$.

11.   A computer program recorded on a computer-readable medium, said program causing a computer to determine foreground and background constraints for a video stream by performing the steps of:

5   receiving foreground and background constraints for a video frame;

generating a matte for the video frame using the foreground and background constraints;

generating foreground constraints for a subsequent video frame using the matte; and

10   generating background constraints for the subsequent video frame using the matte.

12.   The computer program of claim 11, wherein the foreground and background constraints for the video frame are received via user input.

15

13.   The computer program of claim 12, wherein generating foreground constraints for the subsequent video frame includes:

performing a thresholding operation on the matte to generate a binary field.

20   14.   The computer program of claim 13, wherein generating foreground constraints for the subsequent video frame includes:

performing a morphological erosion operation on the binary field.

15.   The computer program of claim 11, wherein generating background constraints

25   for the subsequent video frame includes:

generating a background score field using the matte; and

performing a thresholding operation on the background score field to generate a binary field.

16.   The computer program of claim 15, wherein generating background constraints for the subsequent video frame includes:

        performing a morphological erosion operation on the binary field.

5

17.   The computer program of claim 15, wherein the background score field includes background scores for pixels of the video frame, the background score for a pixel being directly related to the pixel's proximity to a foreground pixel.

10   18.   The computer program of claim 15, wherein the background score field includes background scores for pixels of the video frame, the background score for a pixel being inversely related to a corresponding value for the pixel in the matte.

19.   The computer program of claim 15, wherein generating the background score
15   field includes:

        generating a distance transform from the foreground constraints; and

        generating a weighted combination of an inverse of the matte and an inverse of the distance transform.

20   20.   The computer program of claim 19, wherein generating the weighted combination includes weighting the inverse of the matte by a weight w and the inverse of the distance transform by a weight $1-w$, wherein $0.5 \leq w \leq 0.9$.
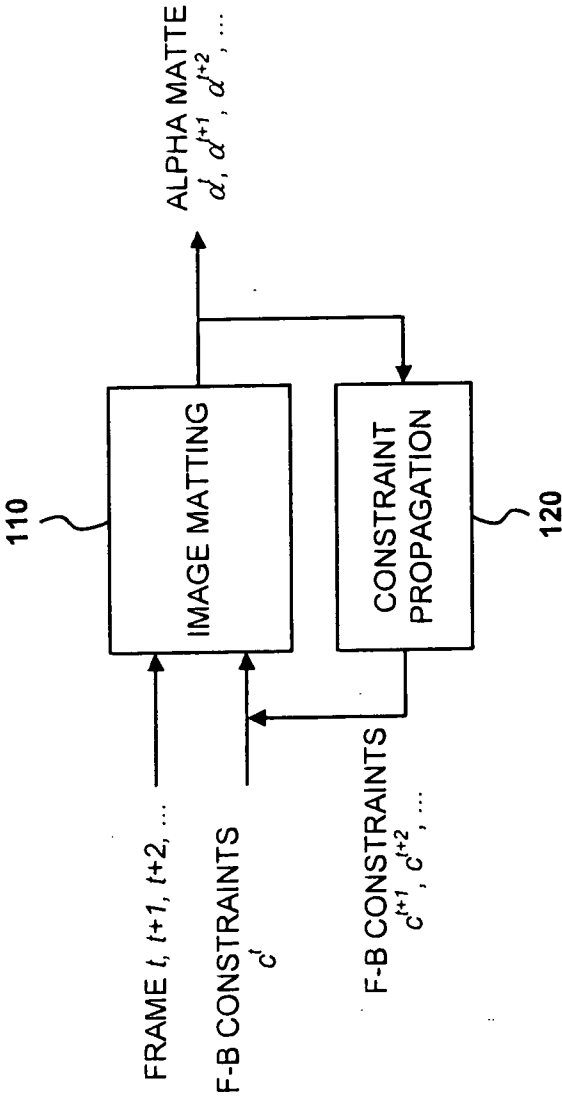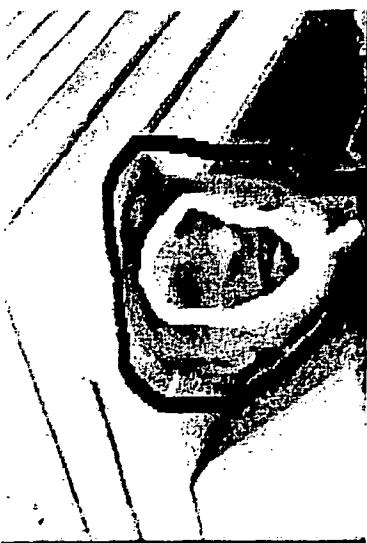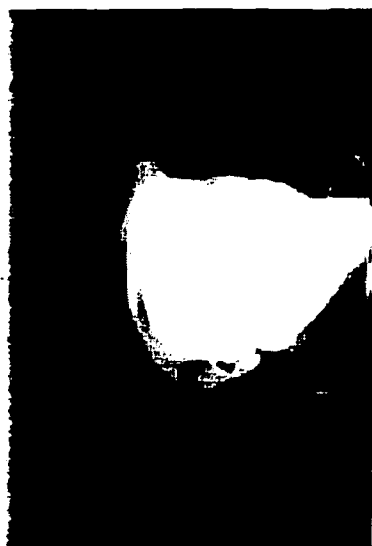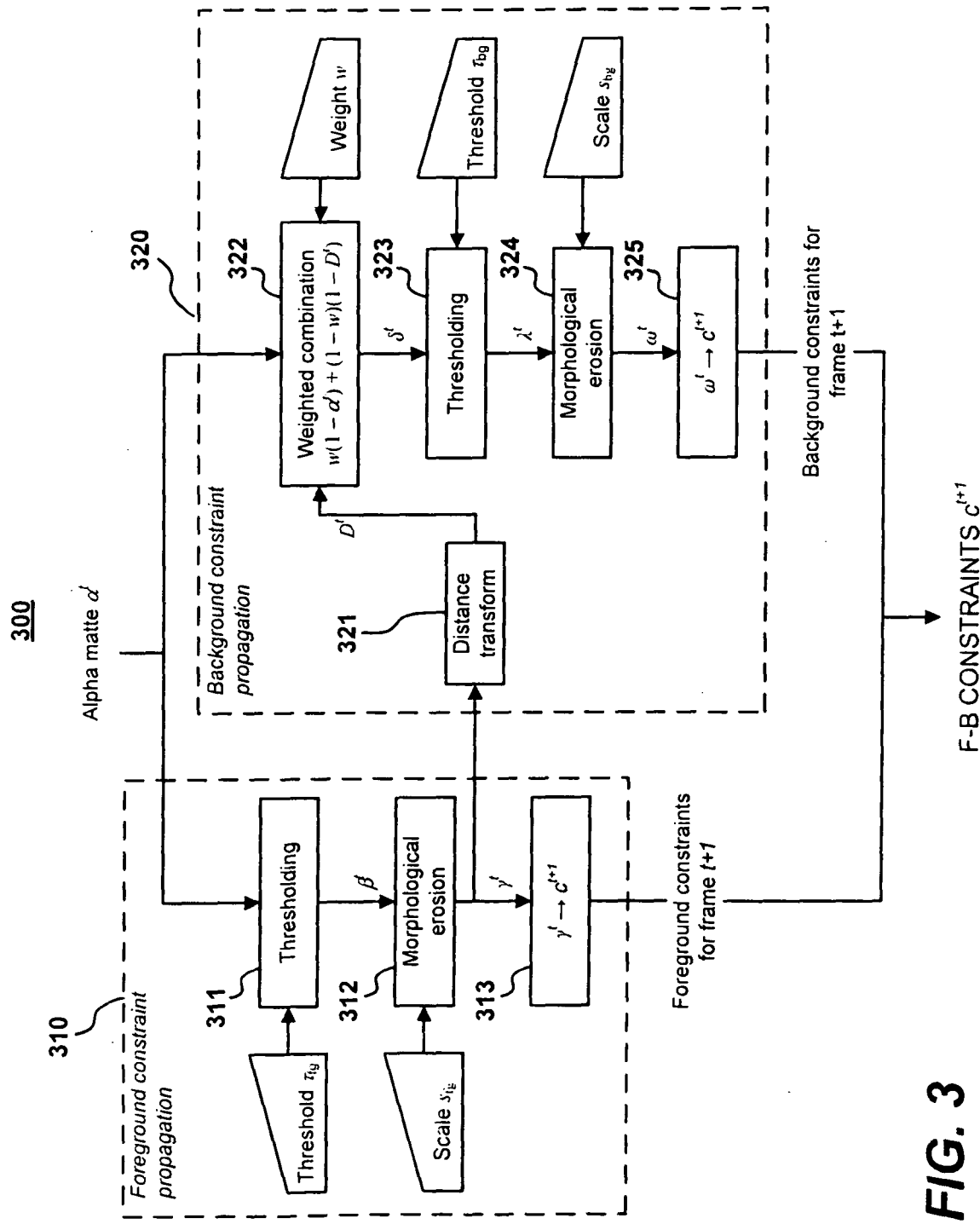
FIG. 1

FIG. 2B

FIG. 2D

FIG. 2A

FIG. 2C

300

Alpha matte α^t

320

Background constraint propagation

Weight w

Threshold τ_{bg}

Scale s_{bg}

322 Weighted combination $w(1-\alpha^t)+(1-w)(1-D^t)$ → δ^t

323 Thresholding → χ^t

324 Morphological erosion → ω^t

325 $\omega^t \rightarrow c^{t+1}$

Background constraints for frame t+1

D^t

321 Distance transform

310

Foreground constraint propagation

311 Thresholding → β^t

312 Morphological erosion → γ^t

313 $\gamma^t \rightarrow c^{t+1}$

Foreground constraints for frame t+1

Threshold τ_{fg}

Scale s_{fg}

F-B CONSTRAINTS c^{t+1}

FIG. 3

FIG. 4A

FIG. 4B

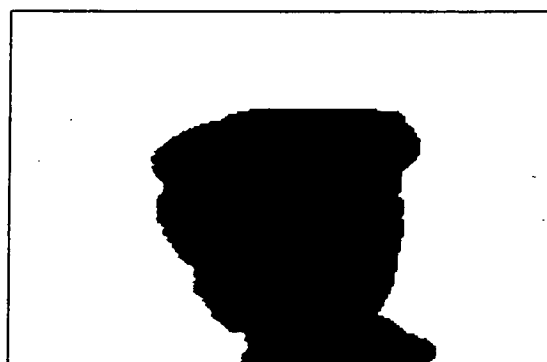**FIG. 4C**

FIG. 4D

FIG. 4E

FIG. 4F

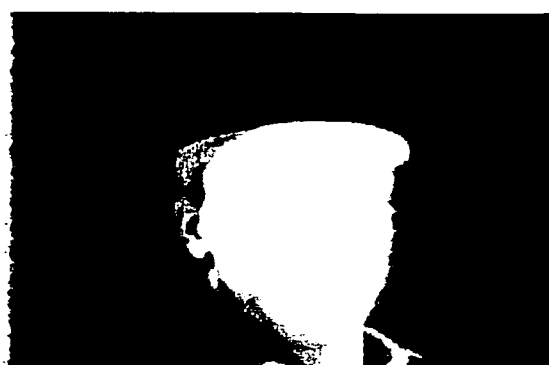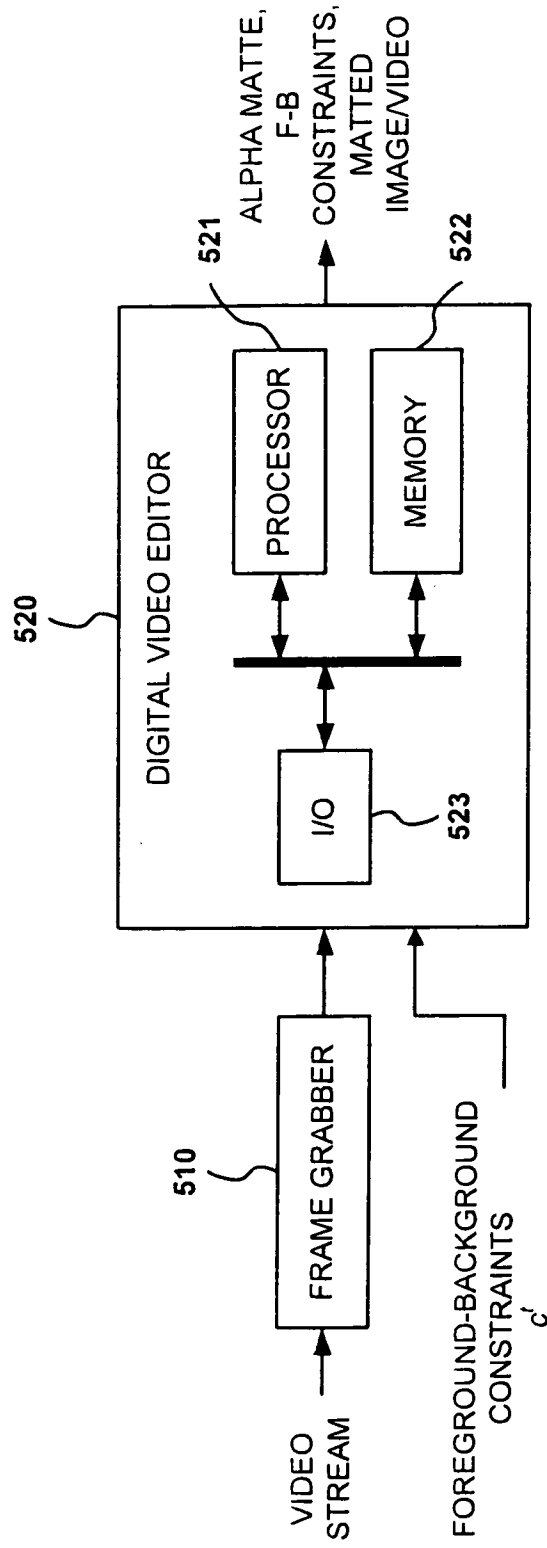*FIG. 5*

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 10/00009

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - G06K 9/36 (2010.01)
USPC - 382/284

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
USPC: 382/284

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
USPC: 382/284

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Electronic Databases Searched: pubWEST(PGPB,USPT,USOC,EPAB,JPAB); GoogleScholar
Search Terms Used: video matting/tooning, morphological erosion/filter/closing, dynamic object tracking, foreground/background
constraints, threshold scaling, distance transform, thresholding, segmenting

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2008/0037872 A1 (LEE et al.) 14 February 2008 (14.02.2008) Entire document, especially: Abstract, para [0003], [0009], [0010], [0013], [0037], [0050], [0054], [0061], [0063]	1-20
Y	US 2007/0297645 A1 (PACE) 27 December 2007 (27.12.2007) Entire document, especially: Abstract, para [0020], [0041]-[0043], [0049], [0088], [0152], [0240], [0241], [0250] and Fig. 2	1-20
A	US 2006/0221248 A1 (McGUIRE et al.) 05 October 2006 (05.10.2006)	1-20
A	US 2006/0039690 A1 (STEINBERG et al.) 23 February 2006 (23.02.2006)	1-20

☐ Further documents are listed in the continuation of Box C.   ☐

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance		
"E"	earlier application or patent but published on or after the international filing date	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed	"&"	document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
16 February 2010 (16.02.2010)	**0 2 MAR 2010**

Name and mailing address of the ISA/US	Authorized officer:
Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450	Lee W. Young
Facsimile No. 571-273-3201	PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (July 2009)