

VIDEO COMPLETION USING HIERARCHICAL MOTION ESTIMATION AND COLOR COMPENSATION

Jin-Hong Kim¹, Rae-Hong Park¹ and Jino Lee¹

¹Department of Electronic Engineering, Sogang University, Seoul, Korea

sosu02kjh@sogang.ac.kr, rhpark@sogang.ac.kr, realltybj@sogang.ac.kr

ABSTRACT

Video completion is used for repairing damaged or missing data in a video sequence. This paper proposes a video completion method using hierarchical motion estimation and color compensation. The proposed method consists of three steps: hierarchical motion estimation, source patch search, and color assignment. If an input video contains brightness change, existing video completion methods give poor results. For overcoming this drawback, we apply two methods: color normalization for motion estimation and source patch search, and color compensation for color assignment. First of all, we construct spatio-temporal Gaussian pyramid for reducing the computation time. We estimate the motion of the input video using color information of each pyramid level. The proposed method uses normalized RGB color in motion estimation. Then, we search the source patch using normalized color information and motion information extracted in the first step. Finally, color compensation using color transform is used for color assignment. Experimental results show that the proposed method gives better visual results than conventional video completion methods.

KEYWORDS

Video completion, motion estimation, Gaussian pyramid

1. INTRODUCTION

Video completion is the method that fills the missing or damaged region in a video. It can be regarded as an extension of image inpainting. Image inpainting fills the missing hole in an image utilizing spatial information [1]–[4]. Image inpainting can be used to repair scratch or remove unpleasant objects in photographs. In addition, it can be applied to restoration of video frame by frame [2]. Nevertheless, image inpainting methods cannot be simply applied to video completion because video has both spatial and temporal information. Video completion methods take account of not only spatial information but also temporal information such as motion to which the human eye is sensitive. Thus, it is important how well we utilize both spatial and temporal information for video completion.

Shen *et al.* used motion manifold and volume rectification for video completion [5]. Motion manifold can reduce video search space from three-dimension to two-dimension. Volume rectification is used for dealing with perspective camera motion. Their method can deal with global illumination change. Patwardhan *et al.* extracted the background and moving foreground from a video sequence using an optical-flow based mask, after which the background and foreground are separately filled [6]. It has a drawback that the video sequence for completion should have the static background, otherwise, it cannot effectively separate the background and foreground. Jia *et al.* proposed a video completion method in which cyclic motion is sampled and intrinsic image separation is used [7]. Their method extracts illumination components from an input video. After separating color and illumination, color and illumination are repaired. Their method can repair the video which has variable illumination such as moving lighting, however,

extracting correct intrinsic images is usually difficult. For repairing moving foreground, damaged moving elements are extracted and repaired by using sampled cyclic motion. Kamel *et al.* applied a Kalman filter to a video completion [8]. Kalman filtering and background subtraction are combined to fill foreground. An exemplar based image completion method used for background repairing. Their method can be applicable to static input video. Shih *et al.* proposed a video completion method for modifying character's actions [9]. They used skeleton based interpolation and exemplar-based image completion for making new actions. Their method can create new actions of human character, but it is hard to create actions of other objects. Lee *et al.* proposed a video completion for multi-view video sequences [10]. When videos have different view point for each frame, simple copy and paste methods like exemplar-based completion method do not give good results. Their method uses corresponding masks for finding matching grid points and transformed matching points. Wang *et al.* used motion states to repair largely occluded moving human [11]. They assumed that periodic human motion can be divided into some motion. Using some categorized motion, human motion can be repaired. Their method is specialized to periodic human motion, thus not used for a video that has non-periodic motion. Chang and Shih proposed a method for repairing old movies [12]. Old movies usually have defects. Firstly, these defects are filled with temporal information. When repairing by temporal information fails, spatial information is used. This method uses temporal and spatial information. Shiratori *et al.* used motion vectors (MVs) obtained by an optical flow technique [13]. This method copies the source MVs that are most similar to the target MVs. Then, it propagates color from neighboring frames using filled MVs. Wexler *et al.* copied pixels from other portions of a video to complete missing regions [14]. The procedure repeats until an objective function value converges. This method works well for static scenes, whereas it doesn't work for dynamic scenes or in the case where there is no periodic motion. Cheung *et al.* proposed a video completion method based on patch-based probability models [15]. They extracted sample patches for unsupervised learning. After unsupervised learning, video epitomes that represent characteristics of input video are constructed. This method can be applicable to other fields of tasks such as image segmentation and super resolution, however it does not give temporal smoothness of the result video well. Fang and Tsai proposed a video completion method for removing advertisement video texts [16]. This method uses a hierarchical model for filling targeted regions. Temporal structure and texture are repaired at the top of the hierarchical model. At the bottom of the hierarchical model, spatial texture region is filled by gradient derivative propagation and duplication. This method considers structure and texture simultaneously, but gradient derivative propagation makes texture blurry. Ding *et al.* proposed a video completion method using rank minimization [17]. Feature descriptors are obtained by rank minimization. Feature descriptors are used for restoring video frames. This method can give good results for static and non-static video, and can handle with non-periodic object motion. Chang *et al.* combined motion compensation and exemplar image completion method [18]. Their method can repair both static and non-static video. If an input video has no temporal continuation, their method shows poor completion results. This paper is an extended version of the paper [19], in which performance comparison of three existing methods is presented for three test images.

The rest of the paper is organized as follows. Section 2 describes in detail the proposed algorithm. First, we explain color optical flow using spatio-temporal Gaussian pyramid. After finding source patch step is described, the source patch is searched hierarchically. Finally, filling color information using color compensation is described. Section 3 gives experimental results and discussions. Finally, Section 4 concludes the paper.

2. PROPOSED VIDEO COMPLETION METHOD USING COLOR COMPENSATION

In this section, we describe the proposed video completion method using color compensation for a video that has brightness change. Figure 1 shows the block diagram of the proposed method. First, we estimate motion of the input color video I using color information. Then, we search

space-time source patches to fill the target region H . Finally, damaged region is filled with color compensated source patch.

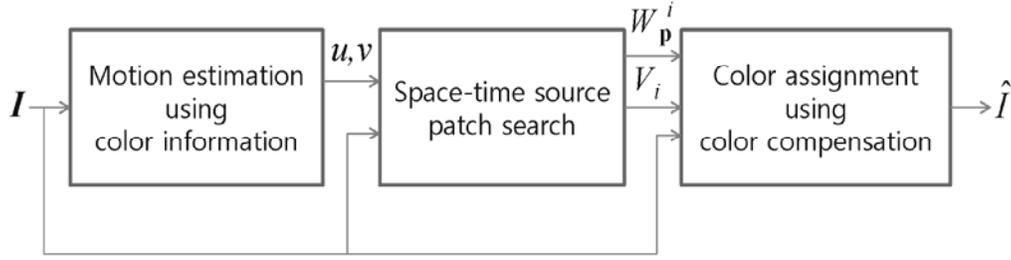


Figure 1. Block diagram of the proposed video completion algorithm

2.1. Hierarchical motion estimation using color information

Temporal information is used to find source patches. We assume that motion information is continuous across the input video. We assign motion information that maintains consistency of motion using the source patches to fill the hole effectively.

Existing intensity-based motion estimation approach extracts inaccurate motion information, because of variation of brightness between patches. Golland and Bruckstein proposed two methods under different assumptions: brightness conservation and color conservation [20]. With two assumptions, we can explain the disadvantage of intensity based motion estimation. Brightness conservation model assumes that there is no brightness change of three color channels between the current and previous frames. Thus, motion information of the current frame can be estimated intensity change of three channels, which is express as

$$\frac{\partial R}{\partial x}u + \frac{\partial R}{\partial y}v + \frac{\partial R}{\partial t} = 0, \quad (1)$$

$$\frac{\partial G}{\partial x}u + \frac{\partial G}{\partial y}v + \frac{\partial G}{\partial t} = 0, \quad (2)$$

$$\frac{\partial B}{\partial x}u + \frac{\partial B}{\partial y}v + \frac{\partial B}{\partial t} = 0 \quad (3)$$

where R , G , and B represent red, green, and blue color components of a color image, u and v are x and y directional motions, respectively. Under the brightness change, intensity variation of each channel affects the accuracy of motion estimation.

Motion estimation under color conservation assumption consists of two equations that use color components, which are expressed as

$$\frac{\partial F_1}{\partial x}u + \frac{\partial F_1}{\partial y}v + \frac{\partial F_1}{\partial t} = 0, \quad (4)$$

$$\frac{\partial F_2}{\partial x}u + \frac{\partial F_2}{\partial y}v + \frac{\partial F_2}{\partial t} = 0 \quad (5)$$

where F_1 and F_2 are color components, and u and v are motions along the x and y directions, respectively. Normalized RGB color is used to remove brightness component and to extract more accurate motion information, which is represented as

$$\begin{pmatrix} r \\ g \\ b \end{pmatrix} = \begin{pmatrix} \frac{R}{R+G+B} \\ \frac{G}{R+G+B} \\ \frac{B}{R+G+B} \end{pmatrix} \quad (6)$$

The normalized RGB represents RGB color components that do not have brightness component. In other words, (6) removes brightness component from RGB color. The normalized RGB color helps extract motion information more accurately than RGB color when an input video has brightness change [12]. Note that F_1 is r and F_2 is g of I in the proposed method, and (4) and (5) are rewritten as

$$\frac{\partial r}{\partial x}u + \frac{\partial r}{\partial y}v + \frac{\partial r}{\partial t} = 0, \quad (7)$$

$$\frac{\partial g}{\partial x}u + \frac{\partial g}{\partial y}v + \frac{\partial g}{\partial t} = 0. \quad (8)$$

Before estimating motion, we construct l -level spatio-temporal Gaussian pyramid to reduce the computation time. A video at coarser level has half resolution of a video at next finer level. Motion estimation is done from coarse to fine level using (7) and (8).

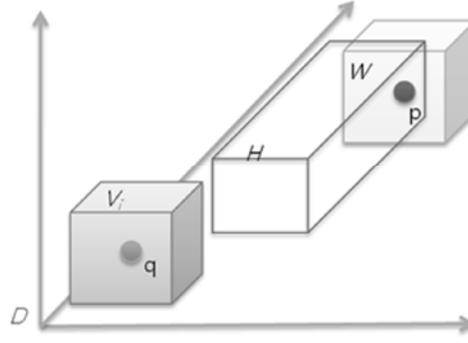
2.2. Space-time source patch search

Damaged region is filled by other portion of an input video using a patch-based filling process. All patches in the target region are compared with patches in the source region and the best matched sample of each target patch is found. We search a minimum difference patch for the patch matching. The proposed method uses normalized RGB color for accurately finding source patches to correct color difference between target and source regions.

We employ the objective function for searching space-time source patch [14]. A completion result \hat{I} can be obtained by maximizing the objective function iteratively, which is defined as

$$Coherence(\hat{I}|D) = \prod_{\mathbf{p} \in \hat{I}} \max_{\mathbf{q} \in D} \text{sim}(W, V_i) \quad (9)$$

where D is a database video that represents portions of I outside of the target region H . Whole video sequence except for target completion region is used as a database video. Figure 2 shows a relation between a target patch W and a source patch V_i between database D and target region H . Space-time pixels \mathbf{p} and \mathbf{q} are denoted by three-dimensional (3-D) vector $(x, y, t)^T$ in each portion of \hat{I} and D , respectively. W is a space-time target patch that includes \mathbf{p} , whereas V_i is a space-time source patch that includes \mathbf{q} . The size of a space-time patch is defined as $(2S+1) \times (2S+1) \times (2S+1)$. When the objective function converges, pixels in the target region can be regarded as correctly repaired.


 Figure 2. Source patch V_i and target patch W in (2).

A global objective function defines global coherence between \hat{I} and D as a local similarity between space-time patches. An important thing in defining a global objective function is how to define a local similarity measure. Image completion usually uses the sum of squared difference (SSD) of color components for a similarity measure. It is not sufficient to use color components alone for a similarity measure in video completion. Because human eyes are more sensitive to motion than to color information, motion information is added for measuring local similarity [14]. Considering this characteristics of human eyes, the local similarity between space-time patches is expressed as

$$\text{sim}(W, V_i) = \exp\left(-\frac{d(W, V_i)}{2\sigma^2}\right) \quad (10)$$

where $d(\cdot, \cdot)$ is the SSD of the five-dimensional (5-D) feature vector between space-time patches, which is expressed as

$$d(W, V_i) = \sum_{\substack{\mathbf{s}^W \in W-H, \\ \mathbf{s}^V \in L}} \|\mathbf{s}^W(x, y, t) - \mathbf{s}^V(x, y, t)\|^2 \quad (11)$$

where L is a set of space-time pixels in V_i . The feature vector \mathbf{s} is defined as

$$\mathbf{s} \equiv (r, g, b, \lambda u, \lambda v)^T \quad (12)$$

where r , g and b are the normalized R , G , and B components, u and v are x and y directional motions, respectively, and λ is a scale factor. For reducing the overall error, σ is set to 75 percent of all distances in current search. Using R , G , and B components for a 5-D feature vector directly, an optimal space-time source patch is not found correctly. Since the normalized RGB color does not contain brightness component, the local similarity using the normalized RGB color model can accurately measure the similarity between space-time patches.

Using spatio-temporal Gaussian pyramid, the computation time of searching source patch can be reduced. After determining the position of source patches at the coarser level, the position of source patches is found at the finer level using the coarse position of source patches.

2.3. Color compensation and color assignment

Brightness changes over the captured images in the same scene with varying viewpoints may create large color artifact when they are merged [21]–[23]. For color constancy of merging image, color difference of overlapped regions must be corrected. Similarly, without the consideration of the difference of brightness between two patches, color information cannot be used directly for color assignment. Assigning colors in the source frame to the target frame without color correction process introduces color distortion since color varies over the input video sequence.

Hence, we compensate for color information of source patches using color transform function. In the proposed method, we use the affine transform model [21]. The transform matrix M between target frame and source frame is defined as

$$M = (Q^T Q)^{-1} Q^T P \quad (13)$$

where P and Q are $N \times 3$ RGB color matrices that are represented as

$$P = \begin{bmatrix} R(\mathbf{p}_1) & G(\mathbf{p}_1) & B(\mathbf{p}_1) \\ \vdots & \vdots & \vdots \\ R(\mathbf{p}_N) & G(\mathbf{p}_N) & B(\mathbf{p}_N) \end{bmatrix} \quad (14)$$

$$Q = \begin{bmatrix} R(\mathbf{q}_1) & G(\mathbf{q}_1) & B(\mathbf{q}_1) \\ \vdots & \vdots & \vdots \\ R(\mathbf{q}_N) & G(\mathbf{q}_N) & B(\mathbf{q}_N) \end{bmatrix}. \quad (15)$$

P and Q contain space-time pixels \mathbf{p} and \mathbf{q} , respectively. N is the number of pixels in the space-time patches. The offset vector of the affine transform is obtained using the mean vectors of pixels in patches \mathbf{m}_p and \mathbf{m}_q , respectively, which are represented as

$$\mathbf{m}_p = [\mu_R^p \quad \mu_G^p \quad \mu_B^p]^T \quad (16)$$

$$\mathbf{m}_q = [\mu_R^q \quad \mu_G^q \quad \mu_B^q]^T \quad (17)$$

where μ_R^p , μ_G^p , and μ_B^p are mean values of R , G , and B values in the target patch, respectively, and μ_R^q , μ_G^q , and μ_B^q are mean values of R , G , and B values in the source patch, respectively. In this way, the transform of color \mathbf{c} at each space-time pixel can be represented as

$$\mathbf{c}(\mathbf{p}) = \mathbf{c}(\mathbf{q})M + \mathbf{m}_p - M\mathbf{m}_q. \quad (18)$$

Color values in whole patches including \mathbf{p} are corrected. Finally, color \mathbf{c} at \mathbf{p} is assigned as

$$\hat{\mathbf{c}} = \frac{\sum_i w_p^i \mathbf{c}^i}{\sum_i w_p^i} \quad (19)$$

where w_p^i is the reliability of i th target patch containing \mathbf{p} , which can be measured using (10) as $\text{sim}(W_p^i, V_i)$, with the source patch most similar to i th target patch.

3. EXPERIMENTAL RESULTS AND DISCUSSIONS

Experimental results show the comparison of video completion performance of the proposed method and three conventional methods: Patwardhan *et al.*'s method [6], Shiratori *et al.*'s method [13], and Wexler *et al.*'s method [14].

In Figure 3, we use the “Walking Woman” video sequence (320×240) with arbitrary brightness change. Figure 3(a) shows example images of synthesized input video that has brightness change. The next frame of the input video is generated, with 10 percent brighter than the previous frame. A damaged frame is created by rectangular mask (50×70). Figures 3(b), 3(c), and 3(d) show enlarged result frames of Patwardhan *et al.*'s method, Shiratori *et al.*'s method, and Wexler *et al.*'s method, respectively. The error propagation can be found in result images of conventional methods. Artifacts are propagated from boundary to inside of the target region in Figures 3(b) and 3(d). These methods fail to reconstruct the shape of a woman's body. Shiratori *et al.*'s method gives better result than other conventional methods, as shown in Figure 3(c). However, artifacts are also founded in the result image. Their method cannot assign proper color information to the target region. In Figure 3(e), on the contrary, the result image of the proposed method shows plausible visual quality.

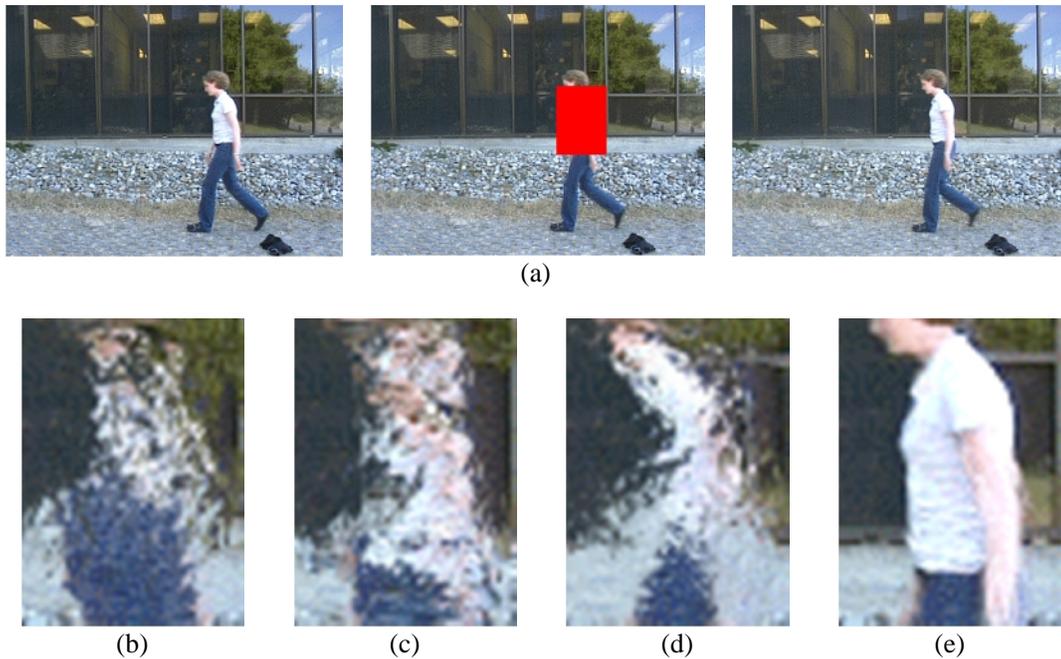


Figure 3. Comparison of result images (320×240, “Walking Woman”). (a) input video, (b) enlarged result frame by [6], (c) enlarged result frame by [13], (d) enlarged result frame by [14], (e) enlarged result frame by the proposed method.

Figure 4 shows the experimental result using a “Chair” video sequence (320×240) used in Jia *et al.*'s experiment [7]. Figure 4(a) shows example images of input video that has brightness change. Figures 4(b), 4(c), and 4(d) show enlarged result frames of Patwardhan *et al.*'s method, Shiratori *et al.*'s method, and Wexler *et al.*'s method, respectively. A result of Patwardhan *et al.*'s method shows boundary distortion. It fails to reconstruct shape of the chair at an intersection point of two edges. Though, the performance of Patwardhan *et al.*'s method is better than that of Shiratori *et al.*'s and Wexler *et al.*'s methods in flat region. Shiratori *et al.*'s results show well reconstructed

contour. However, there are artifacts in the flat region. In Figure 4(c), unnatural edges are found at top of a chair. Wexler *et al.*'s result shows better result than Patwardhan *et al.*'s and Shiratori *et al.*'s methods. However, Figure 4(d) shows regions filled with unnatural colors. The result image of the proposed method shows plausible visual quality as shown in Figure 4(e). It shows sharper edges and more uniform colors than other methods in the completion region.

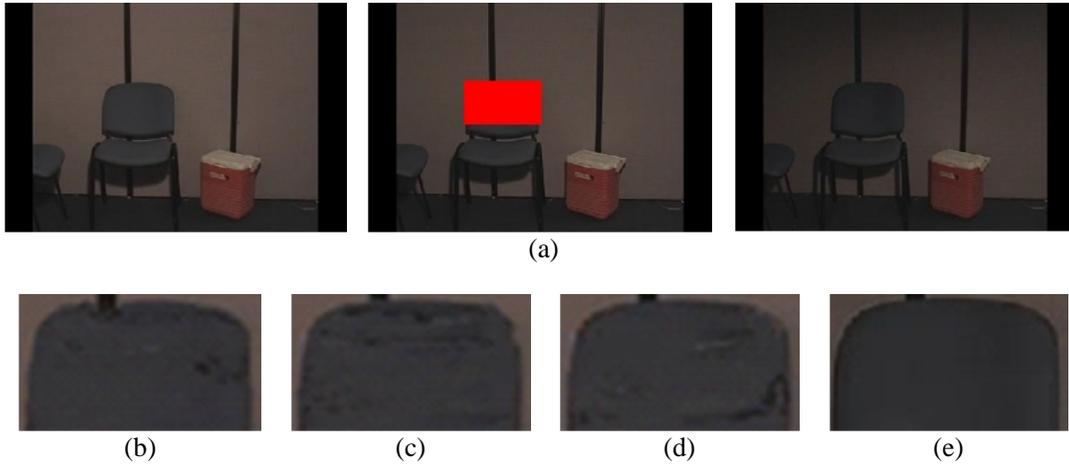


Figure 4. Comparison of result images (320×240, “Chair). (a) input video, (b) enlarged result frame by [6], (c) enlarged result frame by [13], (d) enlarged result frame by [14], (e) enlarged result frame by the proposed method.

Figure 5 shows the real “Door” video sequence (600 × 317) with arbitrary brightness change. Figure 5(a) shows three frames of the input video, in which the damaged region to be completed is indicated by the 50 × 40 rectangle. Figures 5(b), 5(c), 5(d), and 5(e) show enlarged frames reconstructed by Patwardhan *et al.*'s method, Shiratori *et al.*'s method, Wexler *et al.*'s method and the proposed method, respectively. The conventional methods do not give significant color artifacts compared with other test videos, because videos consist of frames with relatively neutral color such as gray. The conventional methods show cloth boundaries not well reconstructed whereas the proposed method gives well reconstructed cloth boundaries.

Result images of conventional methods show the poor performance under brightness change. However, since the proposed method is based on Wexler *et al.*'s method, the computational complexity is almost the same as theirs and the proposed method is applicable to the video sequence with periodic motion.

Table 1 shows performance comparison of four image completion methods in terms of the SSD. The proposed method not only reconstructs smoothly images but also gives the lower SSD values than the conventional methods.

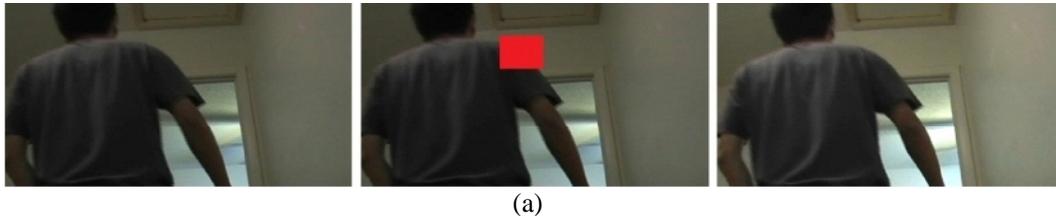




Figure 5. Comparison of result images (600×317, “Door”). (a) input video, (b) enlarged result frame by [6], (c) enlarged result frame by [13], (d) enlarged result frame by [14], (e) enlarged result frame by the proposed method.

Table 1. Performance comparison of four completion methods in term of the SSD.

	Wexler <i>et al.</i> 's method	Shiratori <i>et al.</i> 's method	Patwardhan <i>et al.</i> 's method	Proposed method
Walking Woman	3.62×10^5	3.84×10^5	3.56×10^5	0.86×10^5
Chair	8.11×10^4	8.96×10^4	8.39×10^4	5.32×10^4
Door	3.75×10^4	3.84×10^4	3.61×10^4	1.62×10^4

4. Conclusion

In this paper, we propose an improved video completion method using hierarchical motion estimation and color compensation under brightness change. To extract more accurate motion information, color motion estimation with normalized color information is used. For better similarity measure, we use normalized color information and hierarchical motion information. Because the color information of each frame is different, color compensation using the affine transform is used for color assignment. Experimental results show the effectiveness of the proposed method under brightness change.

REFERENCES

- [1] A. Criminisi, P. Perez, and K. Toyama, “Region filling and object removal by exemplar-based image inpainting,” *IEEE Trans. Image Processing*, vol. 13, no. 9, pp. 1200–1212, Sept. 2004.
- [2] M. Bertalmio, L. Vese, G. Sapiro, and S. Osher, “Simultaneous structure and texture image inpainting,” *IEEE Trans. Image Processing*, vol. 12, no. 8, pp. 882–889, Aug. 2003.
- [3] A. Levin, A. Zomet, and Y. Weiss, “Learning how to inpaint from global image statistics,” in *Proc. Int. Conf. Computer Vision*, vol. 1, pp. 305–312, Nice, France, Oct. 2003.
- [4] S. Yoo and R.-H. Park, “Red-eye detection and correction using inpainting in digital photographs,” *IEEE Trans. Consumer Electronics*, vol. 55, no. 3, pp. 1006–1014, Aug. 2009.
- [5] Y. Shen, F. Lu, X. Cao, and H. Foroosh, “Video completion for perspective camera under constrained motion,” in *Proc. Int. Conf. Pattern Recognition*, vol. 3, pp. 63–66, Hong Kong, China, Aug. 2006.
- [6] K. A. Patwardhan, G. Sapiro, and M. Bertalmio, “Video inpainting under constrained camera motion,” *IEEE Trans. Image Processing*, vol. 16, no. 2, pp. 545–553, Feb. 2007.

- [7] J. Jia, Y.-W. Tai, T.-P. Wu, and C.-K. Tang, "Video repairing under variable illumination using cyclic motions," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 28, no. 5, pp. 832–839, Mar. 2006.
- [8] S. Kamel, H. Ebrahimnezhad, and A. Ebrahimi, "Moving object removal in video sequence and background restoration using Kalman filter," in *Proc. Int. Symp. Telecommunications*, Tehran, Iran, pp. 580–585, Aug. 2008.
- [9] T. K. Shih, N. C. Tan, J. C. Tsai, and H.-Y. Zhong, "Video falsifying by motion interpolation and inpainting," in *Proc. Computer Vision and Pattern Recognition*, 10.1109/CVPR.2008.4587701, Anchorage, Alaska, Jun. 2008.
- [10] S.-Y. Lee, J.-H. Heu, C.-S. Kim, and S.-U. Lee, "An object inpainting algorithm for multi-view video sequences," in *Proc. Int. Conf. Image Processing*, pp. 533–536, San Diego, CA, Oct. 2008.
- [11] H. Wang, H. Li, and B. Li, "Video inpainting for largely occluded moving human," in *Proc. Int. Conf. Multimedia Expo*, pp. 1719–1722, Beijing, China, Jul. 2007.
- [12] R.-C. Chang and T. K. Shih, "Spatial-temporal method for image restoration in aged motion pictures," in *Proc. Int. Conf. Advanced Information Networking and Applications*, pp. 724–729, Washington, DC, May 2007.
- [13] T. Shiratori, Y. Matsushita, X. Tang, and S. B. Kang, "Video completion by motion field transfer," in *Proc. Computer Vision and Pattern Recognition*, vol. 1, pp. 411–418, New York, Jun. 2006.
- [14] Y. Wexler, E. Shechtman, and M. Irani, "Space-time video completion," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 29, no. 3, pp. 463–476, Mar. 2007.
- [15] V. Cheung, B. J. Frey, and N. Jojic, "Video epitomes," *International Journal of Computer Vision*, vol. 76, no. 2, pp. 141–152, Feb. 2008.
- [16] C.-L. Fang and T.-H. Tsai, "Advertisement video completion using hierarchical model," in *Proc. Int. Conf. Multimedia Expo*, pp. 1557–1560, Hannover, Germany, Jun. 2008.
- [17] T. Ding, M. Sznajder, and O. I. Camps, "A rank minimization approach to video inpainting," in *Proc. Int. Conf. Computer Vision*, 10.1109/ICCV.2007.4408932, Rio de Janeiro, Brazil, Oct. 2007.
- [18] R.-C. Chang, N. C. Tang, and C. C. Chao, "Application of inpainting technology to video restoration," in *Proc. Int. Conf. Ubi-Media Computing*, pp. 359–364, Lanzhou, China, Jul. 2008.
- [19] J.-H. Kim, R.-H. Park, and J. Lee, "Video completion using robust motion estimation and color compensation," in *2011 Digest of Technical Papers Int. Conf. Consumer Electronics*, pp. 535–536, Las Vegas, NV, Jan. 2011.
- [20] P. Golland and A. M. Bruckstein, "Motion from color," *Computer Vision and Image Understanding*, vol. 68, no. 3, pp. 346–362, Dec. 1997.
- [21] G. Y. Tian, D. Gledhill, D. Taylor, and D. Clarke, "Colour correction for panoramic imaging," in *Proc. Int. Conf. Information Visualisation*, pp. 483–488, London, UK, Jul. 2002.
- [22] Y. Xiong and K. Pulli, "Fast panorama stitching for high-quality panoramic images on mobile phones," *IEEE Trans. Consumer Electronics*, vol. 56, no. 2, pp. 298–306, May 2010.
- [23] S.-J. Ha, H.-I. Koo, S. H. Lee, N.-I. Cho, and S.-K. Kim, "Panorama mosaic optimization for mobile camera systems," *IEEE Trans. Consumer Electronics*, vol. 53, no. 4, pp. 1217–1225, Nov. 2007.

Authors

Jin-Hong Kim received the B.S. degree in electronic engineering from Sogang University in 2004. Currently he is working toward the M.S. degree in electronic engineering from Sogang University. His current research interests are image inpainting and image enhancement.

Rae-Hong Park was born in Seoul, Korea, in 1954. He received the B.S. and M.S. degrees in electronics engineering from Seoul National University, Seoul, Korea, in 1976 and 1979, respectively, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 1981 and 1984, respectively. In 1984, he joined the faculty of the Department of Electronic Engineering, School of Engineering, Sogang University, Seoul, Korea, where he is currently a Professor. In 1990, he spent his sabbatical year as a Visiting Associate Professor with the Computer Vision Laboratory, Center for Automation Research, University of Maryland at College Park. In 2001 and 2004, he spent sabbatical semesters at Digital Media Research and Development Center, Samsung Electronics Co., Ltd. (DTV image/video enhancement). His current research interests are video communication, computer vision, and pattern recognition. He served as Editor for the Korea Institute of Telematics and Electronics (KITE) Journal of Electronics Engineering from 1995 to 1996. Dr. Park was the recipient of a 1990 Post-Doctoral Fellowship presented by the Korea Science and Engineering Foundation (KOSEF), the 1987 Academic Award presented by the KITE, and the 2000 Haedong Paper Award presented by the Institute of Electronics Engineers of Korea (IEEK), the 1997 First Sogang Academic Award, and the 1999 Professor Achievement Excellence Award presented by Sogang University.

Jino Lee received the B.S. degree in electronic engineering from Sogang University in 2010. Currently he is working toward the M.S. degree in electronic engineering from Sogang University. His current research interests are image inpainting and image enhancement