時空間画像解析による透過画像と反射画像の分離手法

ThandaOo† 川崎 洋† 大沢 裕†

† 埼玉大学工学部情報システム工学科 〒338-8570 さいたま市桜区下大久保255
E-mail: †{thanda,kawasaki,ohsawa}@mm.ics.saitama-u.ac.jp

あらまし 実世界においては、窓ガラスや写真フレーム等のように反射と透過の組み合わさった物体が数多く存在している。これらはガラスの内側と外側の光源環境の違いや、見る方向やガラス材質等の違いにより、見え方が大きく変化するため、画像処理やテクスチャ獲得などにおいて大きな問題となっていた。これまで、反射と透過を分離する研究は数多く行われてきたが、屋外などのような複雑なシーンを撮影したカラー画像を、反射画像と透過画像に分離することは容易ではなかった。本論文では、等速直線運動を仮定した映像を入力とすることで、EPI解析と呼ばれる時空間画像解析手法を用いて、効率よく反射と透過を分離する手法を提案する。提案する手法を用いれば、都市などのように窓ガラスが多数あり、反射と透過が複雑に組み合わさったカラー画像を、反射画像と透過画像に分離することが出来る。

Separation of Reflection and Transparency based on Spatio-Temporal Analysis

Thanda OO†, Hiroshi KAWASAKI†, and Yutaka OHSAWA†

† Department of Information and Computer Sciences, Faculty of Engineering, Saitama University,
Shimo-Okubo 255, Sakura-ku, Saitama, 338-8570 Japan
E-mail: †{thanda,kawasaki,ohsawa}@mm.ics.saitama-u.ac.jp

Abstract The effect of reflection and transparency is superimposed in many real world scenes, which is caused by glass-like shiny and transparent materials. The presence of such incidental effect in a captured image has made it difficult to apply computer vision algorithms and has lead to erroneous results. This paper presents an optimal method for the automatic separation of reflection and transparency layers even if the scene is complicated. The method is based on the epipolar plane image (EPI) analysis. The method is not like the conventional edge-based EPI analysis, but instead it is an edge and line-based color analysis of EPI and it can robustly separate two component layers. To demonstrate the effectiveness of our method, we present the results of experiments using synthesized and real scene images including indoor and outdoor scenes, from which we successfully extracted the reflection and transparency component layers from the input image sequences.

1. Introduction

The presence of reflection and transparency in the same image is caused by glass-like shiny and transparent materials and usually occurs in many natural scenes. The observed color of such a scene is a combination of the light transmitted from an actual object behind the glass and a reflected object (virtual object) in front of the glass. This combination causes errors in some computer vision algorithms, such as the algorithms for object tracking. In addition, this effect strongly disturbs the texture acquisition of a real-world scene, which is a critical research area in computer vision and computer graphics, including 3D city modeling projects. One possible solution to this problem is to separate the component images. Many researchers have tried to separate the reflection component and many valuable methods have already been proposed. However, most proposals have not considered for outdoor scenes. In this paper, we propose a new method to separate the reflection transparency layers which have been caused by an object placed in front of glass and behind the glass, even if the objects have a view-dependent effect such as specularity. This method is based on the epipolar plane image (EPI) analysis. Unlike conventional EPI analysis, which usually analyzed the edges, we propose edge and line-based color analysis of EPI, which can robustly detect the boundary lines of EPI-strips. After detecting the boundary, we can apply a separation algorithm to separate the reflection and transparency components, even in the presence of specularities. The remainder of this paper is organized as follows. In Section 2, we discuss related work. A detailed explanation of EPI based edge and color analysis is described in Section 3. Section 4 introduces a new method to separate the reflection and transparency components which we have proposed. Experiments and results can be seen in Section 5. Finally, in Section 6, we offer our conclusion.

2. Related Works

There has been much research conducted to separate reflection and transparency layers from an input image in various ways. Some proposed methods based on layer motion [14] [13]. Szeliski [13] proposed layer extraction technique even in the presence of reflection and transparency by estimating layer motion based on constrained least square method. Likewise, some other technique [14] [5] based on motion for image enhancement and transparency separation. Schechner worked out to separate transparent layers using focus [7] [9]. Moreover, they have been proposed to separate these real and virtual objects using an optical property called polarization [8], [11], [10]. However, the polarization based separation techniques have a restriction on one assumption: the angle of incidence. Furthermore, this method need a polarization filter to be operate with the camera and need to capture more than one image by rotating the polarization filter for every scene and difficult to apply real-world scenes such as urban images. On the other hand, a number of research works using independent component analysis [4] [2] [3] and layer information exchange [6] have been proposed .

However, most of these works do not consider for outdoor scene and therefore their techniques are not sufficient for practical use. In this paper, we propose a method for the separation of reflection and transparency under a more general condition which allows a view-dependent effect such as specularity and depth disparity. In this method, we adopt an EPI analysis. Originally, EPI analysis was for the acquisition of 3D data from an image sequence [1]. Currently, some research exists which utilizes EPI for color analysis [12]. We also conduct a color-based approach for EPI analysis.

3. EPI Analysis

3.1 Definition of an EPI

In this section, we explain how we create an epipolar plane image and how we perform an EPI analysis based on the nature and geometrical appearance of each EPI strip. The first step is to make a spatio-temporal image volume and slice it horizontally to acquire the EPI. The EPI volume can be constructed by taking a series of images from a moving camera and accumulating these images along the temporal direction, as shown in Fig. 1. The camera's motion is restricted to constant speed and a straight path, and so these restrictions make the analysis easy and robust. Ideally, the frontal surface of any object appears as an area mounted by two distinct parallel boundaries on the EPI. In the rest of this paper, we call this area the EPI strip, or strip. Since we restrict the camera movement along a straight line and the depth d of all the objects are not the same in the real world, all the strips do not lie in a parallel direction. This depth difference gives a special character to the EPI, as shown in Fig.1(bottom).

3.2 Edge-based EPI Analysis

If there is no transparent object in the scene, as in Fig.2(left), we can produce an EPI as in Fig.2(right). While object 2 can be captured for all frames, objects 1 and 3 appear for a short period in the camera view. θ_1, θ_2, θ_3

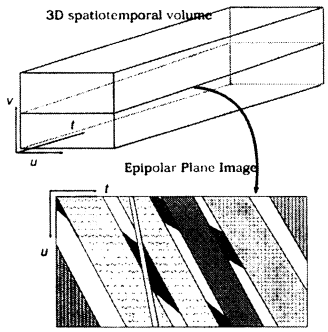


Fig 1 An EPI

in Fig.2 represent the inclination angles of EPI strips 1, 2 and 3, respectively, and we can clearly see that the inclination angles of the EPI strips are different, depending on the depth of the object($\theta_1 < \theta_3 < \theta_2$ with respect to $d_1 < d_3 < d_2$). Furthermore, strip 2 is totally covered by the other two opaque strips at the overlap areas, so we cannot see the object if there is an object in front of it. Therefore, the boundary edge of strip 2 cannot be connected at the overlap area and strip 2 is divided into three separate areas. Since these three areas are separated, we can still understand that these three areas produce an EPI strip by analyzing the edge parallelism and color similarity. With such an edge-based analysis, we can retrieve the 3D information from the EPI and the scene.

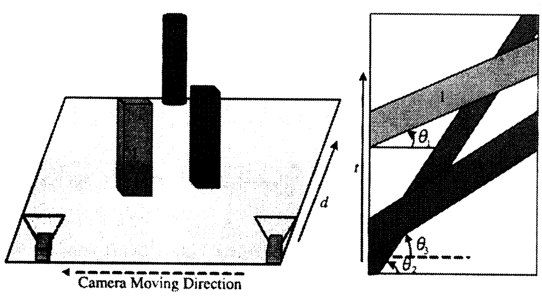


Fig 2 Appearance and nature of actual objects in the EPI

3.3 Color-based EPI Analysis

To conduct the conventional EPI analysis, we have to assume that the object appearance does not change; that is, the object appearance does not depend on the view direction and there is no reflection and transparency. However, in reality, the color sometimes drastically changes depending on the view direction because of the superimposed reflection and transparency and other effects. To conduct a further analysis with an EPI in a real-world scene, we have to understand how color is produced on an EPI and we must include the effects of reflection and transparency. A color change on an EPI can be basically explained by two reasons: one is the changes of material or color of the target objects and the other is the reflection and transparency effect caused by the existing of glass-like object in front of the target scene. It should be noted that we do not consider complicated bi-directional reflectance distribution function (BRDF) in this paper. All colors on an EPI can be explained by the combination of those two effects.

3.3.1 Reflection and Transparency on an EPI

As shown in Fig.3, a reflected object is observed as if it exists on the opposite side of glass; therefore, the object simply describes an EPI-strip on an EPI. However, since glass is transparent, the observed color is a mixture of actual and reflected objects. Therefore, since the reflected object makes a single band, its color changes abruptly when it intersects the EPI-strips of the actual objects and vice versa. Note that, under such conditions, we can still distinguish each EPI-strip robustly; such distinction is usually difficult to achieve by simple image processing techniques such as motion tracking technique applied on the original image sequence.

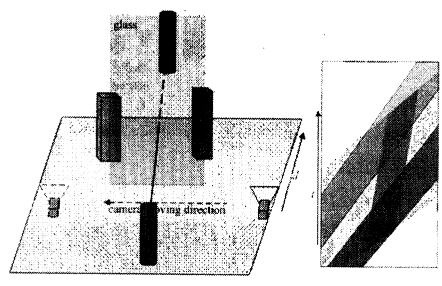


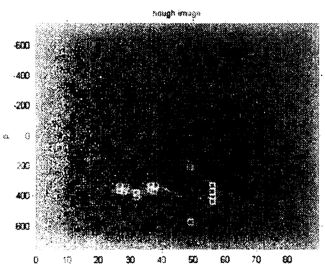
Fig 3 Appearance and nature of transparent and reflected objects in an EPI

4. Separation of Reflection and Transparency

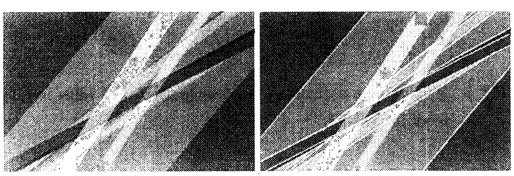
We now describe a technique to separate the two component layers of the EPI and estimate the underlying original colors of the overlap regions. The technique first detects the inclination lines of the EPI-strips. EPI is then rectified by inclination angle of EPI-strip so that trails within strip are vertical. Original color estimation can be done by applying the proposed method along the vertical scan line as describe in Sec 4. 2. Once separation is achieved, the corresponding region is labeled and excluded from further computations.

4. 1 Defining Strips on an EPI

Since the camera is assumed to move linearly, each object in the scene is bounded by two parallel lines on the EPI. As a result, parallel line detection by using Hough transform is sufficient to detect the boundary lines of the EPI-strips as shown in Fig 4(c). The inclination angle of the boundary line in the EPI is directly proportional to the depth of the object. We ordered the EPI-strips based on the inclination angle of the strips by increasing order. Generally the reflected object is assumed far away from the camera, the EPI-strips can be grouped into two according to their inclination angle of boundary lines. The margin of two groups is defined manually. For figure 4(a), EPI-strips those theta values less than fifty ($\theta < 50$) are grouped into one and the rest are another group. Separation method is applied for all strips of a group those having small inclination angle.



(a)



(b) (c)

Fig 4 (a) Hough transform image of (b), white rectangles represent for maximum Hough peaks. (b) The synthesize EPI. (c) The detected boundary lines ($\theta < 50$).

4. 2 Separation of the EPI

Considering the presence of both the reflection and transparency components at the same image point, if we suppose the color of the overlap area is the linear combination of two color components, the observed color of that image point can be described as

$$M_c(x) = f_t \sigma_c^{act}(x) + f_r \sigma_c^{virt}(x) \quad (1)$$

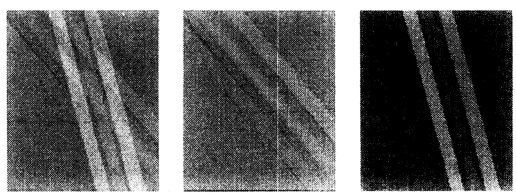
where c represents the type of sensor (r, g and b), $x = \{x, y\}$ is the two-dimensional image coordinate, $\sigma_c^{act}(x)$ and $\sigma_c^{virt}(x)$ represent the color of the actual and virtual objects, respectively. f_t and f_r are the factors of transparency and reflection, respectively. Since all the boundary lines of EPI-strips have already been observed, the decomposition of the input EPI into reflected and transparent layers can be done by the following algorithm. The separation is performed each column of the rectified strip. Since each point in a vertical line of EPI-strip represents for the same image point of the object of each image frame, the minimum color value along the entire line should be an original color of that line due to the linear color combination property of reflection and transparency . The algorithm is as follows.

Algorithm 4..1: SEPARATION(S)

comment: Separation of the EPI S

$(r, c) \leftarrow Size(S)$

$$\textbf{for } i \leftarrow 1 \textbf{ to } c \begin{cases} temp \leftarrow GetColumn(S, i) \\ Orgcolor \leftarrow Min(temp) \\ SetColumn(S, i, Orgcolor) \end{cases}$$



(a) (b) (c)

Fig 5 Separation of the synthesized EPI.(a) input EPI (b),(c) separation results.

Another component image can be obtained by subtracting the result image from the original EPI. The separation results of Fig.5(a) can be seen in Fig.5(b) and (c). The separated results show the algorithm 4..1 is good enough for the separation of EPI, since this EPI is produced from the synthesized image sequence(synthesized-EPI) and the surface of all objects are totally plane (no depth disparity, e.g. . window frames of the building in the outdoor scene). Further, the estimation of minimum color is not suitable for real-EPI because of noise and artifact. Therefore, the histogram thresholding method is used for real-EPI separation instead of minimum color estimation. The basic task of this method takes the pixel value of the first peak nearest to zero intensity and substitute for all pixels which are brighter than that along the entire scan line. Fig 6 shows the three possible patterns of histogram as an example of line-based color analysis. For Fig6(a),(b), there is no ambiguity to select the first peak. Comparing with several experimental results, the first peak of Fig6(c) is small and it should be neglected, because this type of peak can be occurred by noise and artifact. After applying the separation algorithm the EPI-strip is rectify back to the original one. Since there remains base color ambiguity with this algorithm, this technique cannot produce the correct color value. However, the result can be effectively used for most computer vision algorithms. Furthermore, for texture acquisition purposes, human interaction can produce a reasonable result.

4.3 Separation of the Original Image

By using the result of the decomposition of the EPI as described above, we can separate the original image into two component images by two ways. The first is a straightforward method which applies a separation algorithm to each EPI for all horizontal lines (we call this the iterative method), and the second is based on color clustering. However, the iterative method is not practical because the accumulation of small errors in the offset for each frame becomes large and the processing of all EPIs is time-consuming. Therefore, the color clustering method is preferable.

This approach first clusters the input images, and then finds the correspondent color in the EPI for separation. The detailed procedure is as follows.

- create sparse EPIs from the captured image sequence and decompose the EPIs by the separation algorithm.

- get the (u, v) coordinate and the color information (r, g, b) from the EPI for the desired input frame as an initial point.

- perform color clustering of the original image by the region-growing, which starts with an initial point and merges neighboring pixels by using their color information in 3D space.

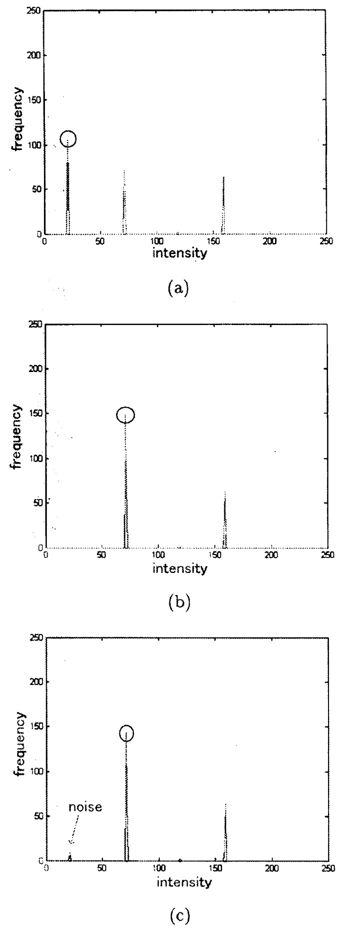


Fig 6 The possible patterns of histogram, black circles show the thresholding results. (a) and (b) show the patterns without noise, (c) shows the noise included pattern.

- after clustering, the component image can be extracted by substituting the expected original color for the clustered pixels, which can be obtained from the resulting EPI.

5. Experiments

We performed several experiments to test the effectiveness of our method. In the following two experiments, we used a synthesized image sequence and a real image captured in our laboratory and outdoor scenes.

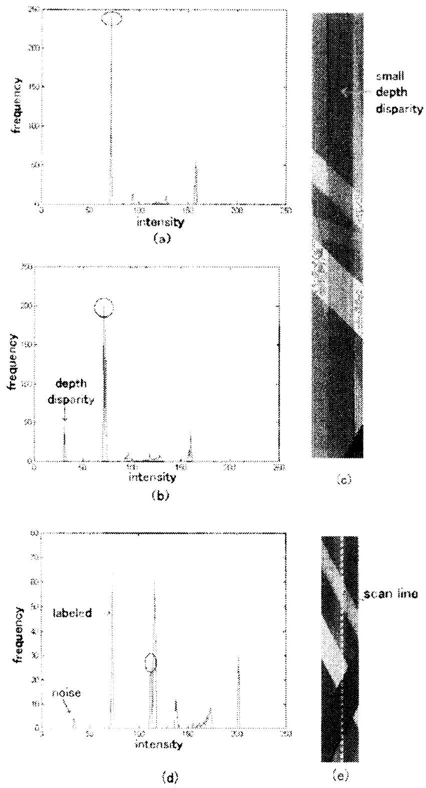


Fig 7 The example histograms of scanned lines, black circles show the thresholding results. (a) and (b) show non-existing and existing depth disparity on scanned line of (c). (d) The thresholding result of scanned line of (e).

5.1 Synthesized images

The image sequences used to test our method have been created by using CG software. As a target object, we constructed a model room which has a front wall covered by glass to create a reflected image of the objects placed in front of the glass wall, as well as transparent images of the objects placed inside the room. We assumed that the factors of reflection and transparency for all captured frames are constant, and the camera motion follows a straight path that produces regularly sampled images for creating the EPI volume. Some histogram thresholding results can be seen

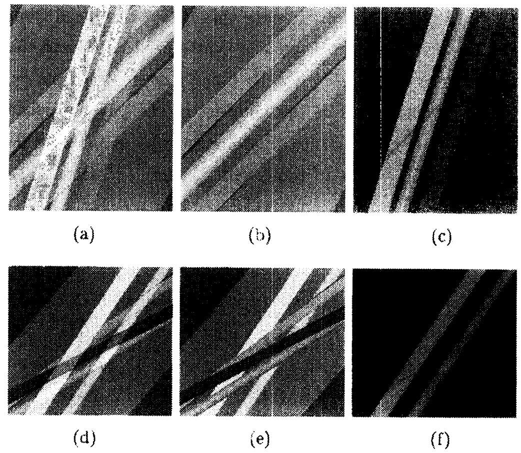


Fig 8 Result of the EPI analysis. (a),(b) Input EPI with view-dependent effect. (c),(d),(e),(f) Separation results

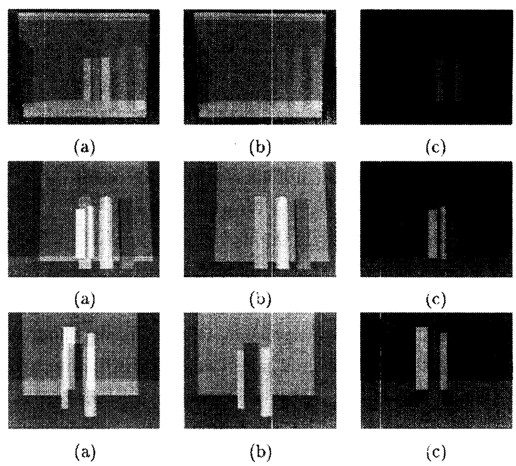


Fig 9 The separation of synthesized images. First column shows the input synthesized images. Separation results can be seen in the second and third columns.

in Fig 7. Fig 7(a) and (b) show the histogram thresholding results of two scanned lines from Fig 7(c). In Fig 7(b), the first peak is not small as noise and it should be the selected peak of that line. However, it is neglected. Because, the thresholding algorithm is not only consider for the noise but also consider for such type of peak, which has been caused by depth disparity. The original color estimation for this type of histogram has been performed by comparing with the thresholding results of successive lines. The selected result proved that the proposed method can overcome the existing of small depth disparities on the surface of the objects. For figure 7(d), the method selects the third peak because the second(red) peak was already labeled in the previous calculation. Fig.8 is the result of EPI analysis. The EPI was successfully separated into component images can be seen in second and third columns. The first row shows that the proposed method can separate two layers even in the presence of specularities in both real and actual objects. However, in Fig.8(c), we can observe small artifacts on EPI strips which are caused by color saturation on the synthesized image. To avoid such artifacts, using a high dynamic range image is a practical solution.

Fig 9 shows the separation results of original images. The left column of Fig.9 shows an arbitrary frame of three input image sequences, and the recovered transparency and reflection component images are shown in the middle and right columns.

5.2 Real Images

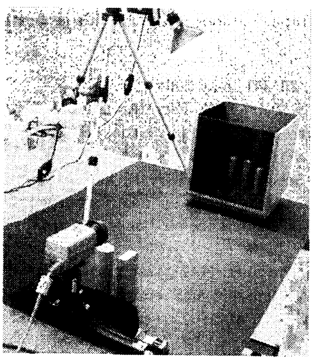
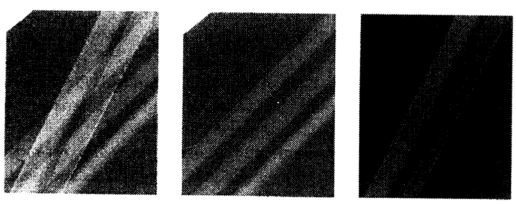


Fig 10 The scene of the real image capturing process.



(a) (b) (c)

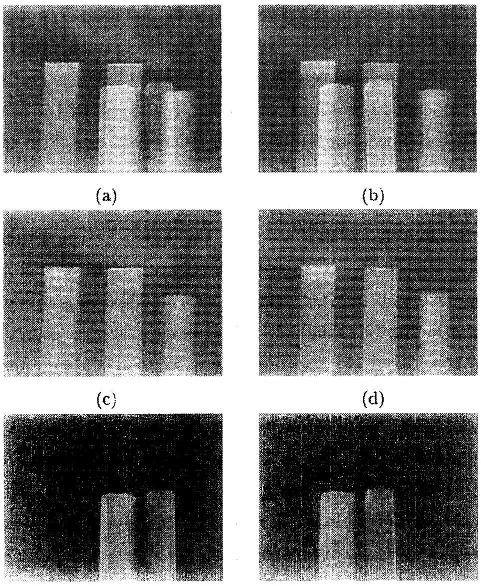
Fig 11 Separation of the real EPI.(a) input EPI (b),(c) separation results.

We have conducted several tests on real images captured using Sony three-CCD(640×480) digital camera. The motorized stage has been used to control the linear movement of the eamera as shown in Fig.10. Fig.11 shows the input EPI and separation result of real image sequence. The original images and their separation results are presented in the Fig.12.

In Fig.11, we can see that the input EPI is successfully separated into component images. However, for the separation of original images, we can observe some artifacts in Fig.12(e) and (f). Main reason for this error can be considered as: the object has weak specularity on it and such view-dependent effects produce color differences in the final result. In addition, we based the method on the assumption which is the color combination of reflection and transparency is linear. As a result, although the proposed method can extract two component layers even in the present of view-dependent effect, the subtracted layer can be affected by some noise or artifact. Therefore, in order to produce a better result without such artifacts, we need further research for color separation method and to reduce assumption for more reliable with natural scene. We have also conducted outdoor image sequences captured by a car-mounted video camera. Fig.13(top) shows the input image and the separation results can be seen in Fig.13(middle and bottom). The color of Fig.13(bottom) is enhanced to make it easier to see. In this case, we used only three sparse EPIs to recover the result images, because the structure of component layers is simple. More EPIs are required to separate complicated scene.

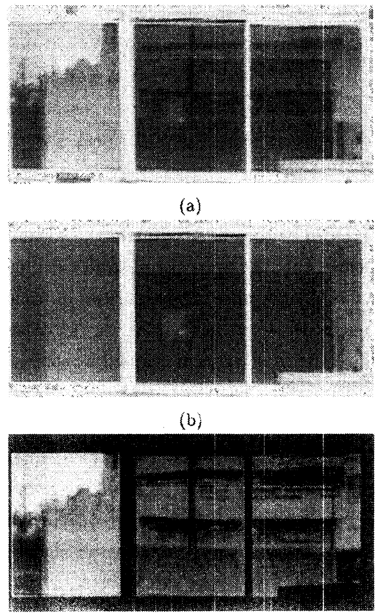
6. Conclusion

In this paper, we proposed a new EPI analysis based on a color analysis, since the conventional EPI analysis does not consider the view-dependent effects of reflection and transparency, which usually exist in real-world scenes. Our proposed method completely assumes these complicated effects and successfully analyzes them. By using our EPI analysis, a scene consisting of glass-like objects which produce both a transparent and reflective effect could be robustly separated into component images. Furthermore, most of our separation method could be performed automatically. Since, many computer vision algorithms usually fail to handle the complicated scene images, our technique can provide



(a) (b) (c) (d) (e) (f)

Fig 12 The separation of real images. First row shows the input real images. Separation results can be seen in the second and third rows.



(a) (b) (c)

Fig 13 Input outdoor image and it's separation results

a practical solution by separating the image into component images. For city modeling purposes, since many buildings are typically covered with glass windows and it is difficult to retrieve textures of good quality, our technique can provide a practical solution.

References

[1] R. Bolles, H. Baker, and D. Marimont. Epipolar plane image analysis:an approach to determining structure from motion. *International Journal of Computer Vision*, 1:7–55, 1987.

[2] A. M. Bronstein, M. M. Bronstein, M. Zibulevsky, and Y. Y. Zeevi. Blind separation on reflections using sparse ICA. In *4th International Symposium of Independent Component Analysis and Blind Signal Separation*, pages 227-232, April 2003.

[3] H. Farid and E. H. Adelson. Separating reflections from images using independent components analysis. *Journal of the Optical Society of America*, 16:2136–2145, 1999.

[4] Hermanto, A. K. Barros, T. Yamamura, and N. Ohnishi. Separating virtual and real objects using independent component analysis. *IEICE TRANS*, E84-D, Sept 2001.

[5] M. Irani and S. Peleg. Motion analysis for image enhancement: resolution, occlusion, and transparency. *Journal of Visual Communication and Image Representation*, 4(4):324–335.

[6] B. Sarel and M. Irani. Separating transparent layers through layer information exchange. In *ECCV*, volume 4, pages 328–341, may 2004.

[7] Y. Y. Schechner, N. Kiryati, and R. Basri. Separation of transparent layers using focus. *International Journal of Computer Vision*, 39:25–39, 2000.

[8] Y. Y. Schechner, N. Kiryati, and J. Shamir. Separation of transparent layers by polarization analysis. In *Scandinavian Conference on Image Analysis*, volume 1, pages 235–242, 1999.

[9] Y. Y. Schechner, N. Kiryati, and J. Shamir. Blind recovery of transparent and semireflected scenes. In *Computer Vision and Pattern Recognition*, volume 1, pages 38–43, 2000.

[10] Y. Y. Schechner and J. Shamir. Vision through semireflecting media: polarization analysis. *Optics Letters*, 24:1088–1090, 1999.

[11] Y. Y. Schechner, J. Shamir, and N. Kiryati. Polarization-based decorrelation of transparent layers: The inclination angle of an invisible surface. In *ICCV*, volume 2, pages 814–819, 1999.

[12] R. Swaminathan, S. B. Kang, and R. Szeliski. On the motion and appearance of specularities in image sequences. In *ECCV*, volume 1, pages 508–523, 2002.

[13] R. Szeliski, S. Avidan, and P. Anandan. Layer extraction from multiple images containing reflections and transparency. In *CVPR*, pages 1246–1253, June 2000.

[14] J. Toro, J. Owens, and R. Medina. Using known motion fields for image separation in transparency. *Pattern Recognition Letters*, 24:594-605, Jan 2003.