Separation of Reflection and Transparency using Epipolar Plane Image analysis

Thanda Oo[1], Hiroshi Kawasaki[1], Yutaka Ohsawa[1], and Katsushi Ikeuchi[2]

[1] Saitama University, Department of Information and Computer Science
255, Shimo-okubo, Sakura-ku, Saitama, 338-8570, Japan
`{thanda, kawasaki,ohsawa}@mm.ics.saitama-u.ac.jp`
[2] The University of Tokyo, Institute of Industrial Science,
6-4-1 Komaba, Meguro-ku, Tokyo, 153-8505, Japan
`ki@cvl.iis.u-tokyo.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. Moreover, it disturbs the texture acquisition of the real-world scene. This paper presents an optimal method for the separation of reflection and transparency components. The method is based on the Epipolar Plane Image (EPI) analysis. The method is not like the ordinary edge-based EPI analysis, but instead it is an edge and color-based EPI analysis. To demonstrate the effectiveness of our method, we present the results of experiments using synthesized and real images which include indoor and outdoor scenes, from which we successfully extracted the reflection and transparency components from the input image sequences. . . .

1 Introduction

Texture acquisition of a real-world scene is one of the critical research areas in computer vision and can be used in other application areas such as computer graphics (CG) including 3D city modeling projects. Moreover, to acquire texture without noise (e.g., a shadow, a specularity, a reflected image) is vital for such work. Generally, many of the buildings are covered with glass windows and glass usually produces reflection and transparency effects. Therefore, 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 situation strongly disturbs the texture acquisition of the real-world scenes. One possible solution to this problem is to separate the component images. Many researchers have tried to separate the reflection and transparency components and many valuable methods have already been proposed. Some proposed methods are based on layer motion [1] [2]. Szeliski [2] proposed layer extraction technique even in the presence of reflection and transparency based on constrained least square method. Likewise, some other techniques [1] [3] based on motion for image enhancement and transparency separation are proposed. Schechner worked out to separate transparent layers using focus [4][5]. Moreover, there have been some proposed methods to separate real and virtual objects using an optical property called polarization [6] [7] [8]. On the other hand, a number of research works using independent component analysis [9] [10] [11] and layer information exchange[12] have been proposed. Some of above mentioned methods need a polarization filter to be operated with camera and need to capture more than one image by rotating the polarization filter or focusing either layer separately for every scene.

According to our knowledge, all of the previously proposed methods are considered only for static camera and/or single depth. Moreover, for the purpose of the texture acquisition of real-world scene, a huge amount of outdoor scene images are originally required and the captured images usually contain 3D objects. As a result, these methods can not be applied for the texture acquisition of outdoor scene. In this paper, we propose a new method to separate the reflection component form image sequence which has been taken by a motion camera. This method is based on the epipolar plane image (EPI) analysis. Unlike previous EPI analysis, which usually analyze the edges, we propose a color-based EPI analysis, which can robustly separate two component layer images.

The remainder of this paper is organized as follows. A detailed explanation of EPI analysis is described in Section 2. In Section 3, we discuss on the separation method of reflection and transparency components. Experimental results can be seen in Section 4 and we provide conclusion in Section 5.

2 EPI Analysis

2.1 EPI construction and conventional EPI analysis

EPI can be produced by accumulating epipolar line in each frame of image sequence along the time axis. The first step is to make spatio-temporal image volume and slice it horizontally to acquire EPI. The camera motion is assumed to constant speed and straight path. Certainly, the restriction is not strictly required in actual experiment, because we can use GPS, gyro sensor and other vehicle speed sensors. The camera is set to arbitrary direction, therefore the rectification of captured image sequence is required before accumulation to make spatio-temporal image volume. Ideally, the frontal surface of any object appears as an area bounded by two distinct parallel boundaries on the EPI (we call this area the EPI-strip, or strip). Since we restrict the camera movement along a straight line and the depth 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). We can clearly see that the inclination angles of the EPI strips are directly proportional to the depth d of the object. Furthermore, strip 2 is totally covered by the other opaque strips at the overlap areas. Therefore, the boundary edges of strip 2 cannot be detected at the overlap areas and strip 2 is divided into separate areas. Since the areas are separated, we can still understand that these areas produce an EPI strip by analyzing the edges parallelism and color similarity. With such an edge-based analysis, we can retrieve the 3D information from the EPI and the scene. Such kind of EPI analysis assumes the object appearance does not depend on the view direction.

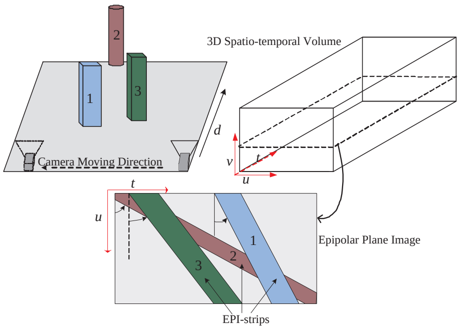


Fig. 1. Appearance and nature of actual objects in the EPI.

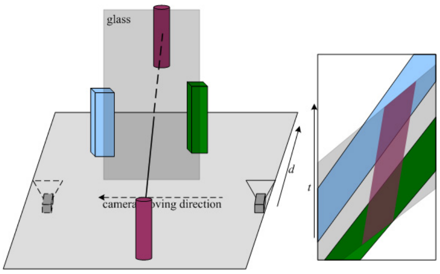


Fig. 2. Appearance and nature of transparent and reflected objects in an EPI.

2.2 Color-based EPI Analysis

In a real-world scene, the color sometimes drastically changes depending on the view direction because of the superimposed reflection and transparency, specularity 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 view dependent effect. We have considered only for superimposed reflection and transparency effect in this work. 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 reflection and transparency effect caused by the glass-like shiny objects in the target scene. It should be noted that we do not consider complicated bi-directional reflectance distribution function (BRDF) in this paper.

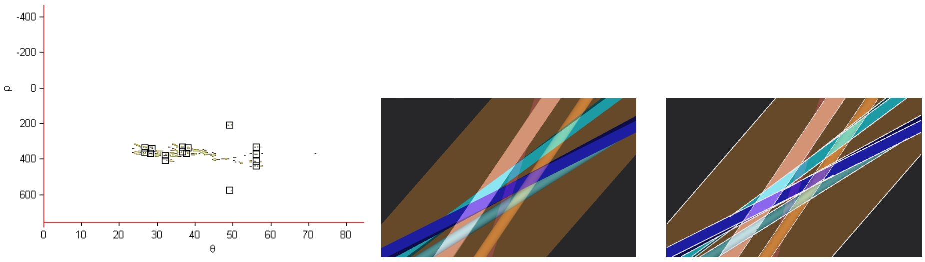
2.3 Reflection and transparency on an EPI

As shown in Fig.2, 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 transparent and reflected objects. Therefore, since the reflected object makes a single band, its color changes abruptly when it intersects the EPI-strips of the transparent 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 tracking applied on the original image sequence.

3 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 3.2. Once separation is achieved, the corresponding region is labeled and excluded from further computations.

3.1 Defining Strips on an EPI



(a) (b) (c)

Fig. 3. (a) Hough transform result of (b), (b) Input EPI, (c) The detected boundary lines.

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. We used only high energy peaks of the Hough space to detect the distinct edges such as boundaries of the building. Fig3(a) represents the selection result of 16 maximum energy peaks of the Hough transformation result of Fig.3(b) and Fig.3(c) shows the detection result. Generally the reflected object (virtual object) is assumed farther than the target objects (wall of the building and other actual objects close to the wall). Therefore, the inclination angles of reflected EPI-strips usually greater than that of the EPI-strips of the actual objects. Then, we detect the boundaries of all EPI-strips and the separation method is applied to each EPI-strip within the detected boundary lines in the increasing order of inclination angles until the whole image area is applied. For Fig.3(b), we could recover all overlap areas by applying the proposed separation method to EPI-strips those inclination angles are lass than 50.

3.2 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

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

where c represents the type of sensor (r, g and b), (x, y) is the two-dimensional image coordinate, $\sigma_c^{act}(x, y)$ and $\sigma_c^{virt}(x, y)$ represent for the color of the actual and virtual objects, respectively. f_t and f_r are the factors of transparency and reflection, respectively. For simplicity, equation (1) can be rewritten as

$$M_c(x, y) = A_c(x, y) + V_c(x, y) \tag{2}$$

The system first rectified the EPI-strip by it's inclination angle, so that trails within the strip are vertical. The separation is performed to each vertical line 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, equation (3) should be true for every pixels within one scan line.

$$A_c(x, y_i) = A_c(x, y_j) \qquad \ldots \qquad (i \neq j) \tag{3}$$

from equation (2) and (3) we can obtain

$$M_c(x, y_i) - M_c(x, y_j) = V_c(x, y_i) - V_c(x, y_j) \tag{4}$$

From this equation, V_c values for all pixels can be calculated if we assume one V_c value as an initiator. Therefore we can obtain infinite sets of V_c values for that line due to initiators. Since all pixels in the scan line are collected from same image point 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. Therefore, $V_c(x, y_i)$ is assumed to 0 where $M_c(x, y_i)$ is the minimum of entire scan line. In our actual implementation we estimate the minimum color along entire scan line and substitute it for all pixels along entire line. However, it is not suitable to apply real-world EPI because of noise and artifact. Therefore, the histogram thresholding method is used for real-EPI separation. The basic task of this method takes the pixel value of the first peak nearest to zero intensity, which is larger than threshold, and substitute for all pixels which are brighter than that along the entire scan line. After applying the separation algorithm the EPI-strip is rectified back to the original one. Another component image can be obtained by subtracting the result image from the original EPI. Since there remains base color ambiguity with this method, this technique cannot produce the correct color value. However, the result can be effectively used for texture acquisition of the real-world scene and human interaction can produce a reasonable result.

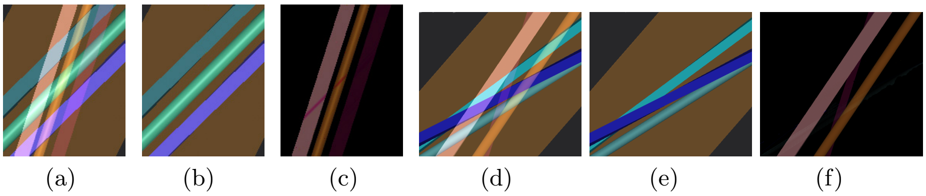

(a) (b) (c) (d) (e) (f)

Fig. 4. Result of the EPI analysis. (a),(d) Input synthesized EPIs. (b),(c),(e),(f) Separation results.

3.3 Separation of the Original Image

By using the decomposition results of the EPIs as described above, we can separate the original image into two component images by two ways. The first is a straightforward method which creates EPIs for all horizontal lines and applies the separation algorithm to each EPI (we call the iterative method), and the second is based on color clustering.

The detailed procedure of second approach is as follows.

- create sparse EPIs from the captured image sequence and decompose the EPIs by the separation algorithm.
- get the (x, y) 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 method, which starts from an initial point and merges neighboring pixels by using their color information in 3D space.
- 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.

4 Experiments

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

4.1 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 along a straight path that produces regularly sampled images for creating the EPI volume. The EPI was successfully separated into component images can be seen in Fig.4. However, in Fig.4(c), we can observe small artifacts on EPI strips which are caused by color saturation on the synthesized images. To avoid such artifacts, using a high dynamic range image is a practical solution.

The left column of Fig.5 shows an arbitrary frame of two input image sequences, and the recovered transparent and reflected component images are shown in the middle and right columns.

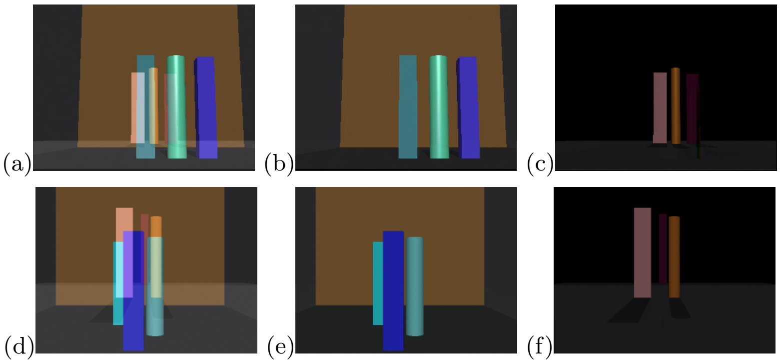


Fig. 5. The separation results of synthesized images. First row: the input synthesized image with specularity and its separation results. Second row: the input synthesized image including different depth objects and separation results.

4.2 Real Images

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 camera in the indoor image capturing process as shown in Fig.6(a). Fig.6 (first and second rows) show the input EPIs and separation results of indoor image sequences. The original images and their separation results are presented in the Fig.7. In Fig.6(d) and Fig.7(c), we can observe some artifacts because of the non-linearity of camera sensor since we can successfully separate synthesized images. Therefore, the linearization of camera sensor is required before the proposed method is applied. For the separation of original indoor image as described in Fig.7(a), we used second approach (color clustering), because the image texture is simple. The another image sequence is decomposed by iterative method as shown in Fig.7 second row.

Car mounted video camera has been used to capture outdoor image sequences by controlling constant car speed and driving along straight path. Since we restrict the constant car speed, a tiny un-constancy of car speed has occurred between each successive image frames. Therefore, the trials within the rectified EPI-strips are not strictly vertical as shown in Fig.8 (a) and the separation results of original image is almost noisy as describe in Fig.9(second row). To

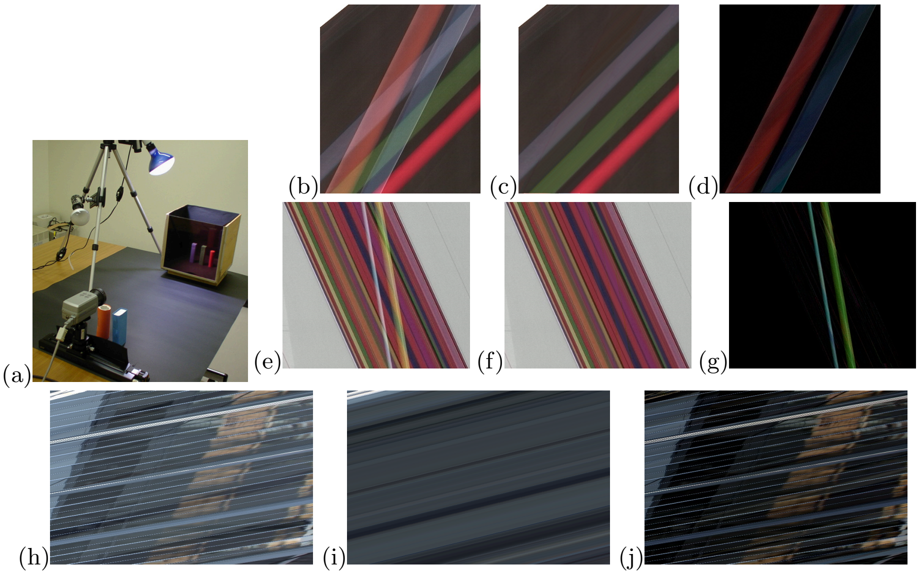


Fig. 6. The scene of the indoor image capturing process (a) and EPI Separation results of real images. (b),(e) input indoor EPIs. (c),(d),(f),(g) separation results of indoor EPIs. (h) input outdoor EPI. (i),(j) separation results.

solve this problem, we first apply simple and straightforward method, which is to reduce the number of image frames in EPI until the rectified strip appears as vertical as shown in Fig.8 (b). Since this method could produce a reasonable result as described in Fig.9 third row, the method will fail when the number of image frames in reduced EPI are too few. The estimation of car speed and adjusting it to every successive image frame is required to produce more modest results. For this purpose, we implement the algorithm, which sampled features from non-transparent image area and tracking for all image frames to detect the motion speed in pixel. The detected pixel difference values have been used to fit the rectified EPI-strip to vertical as described in Fig.8 (c). The separation results of original outdoor image can be seen in Fig.9 bottom row. As a result, the proposed method can extract two component layers even the image texture is complicated as we can observe specular effect on the reflected buildings.

5 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

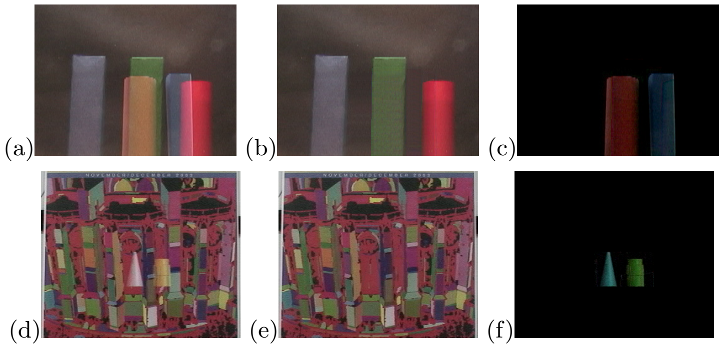
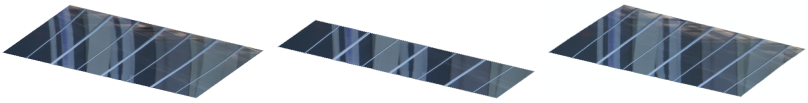


Fig. 7. Separation of indoor images.(a),(d) input images. (b),(c),(e),(f) separation results.



(a) (b) (c)

Fig. 8. Rectified EPIs of outdoor image sequence (a) using all EPI range (b) using reduced EPI range (c) fitted with tracking result.

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 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. Toro, J., Owens, J., Medina, R.: Using known motion fields for image separation in transparency. Pattern Recognition Letters **24** (2003) 594–605
2. Szeliski, R., Avidan, S., Anandan, P.: Layer extraction from multiple images containing reflections and transparency. In: CVPR. (2000) 1246–1253
3. Irani, M., Peleg, S.: Motion analysis for image enhancement: resolution, occlusion, and transparency. Journal of Visual Communication and Image Representation (**4(4)**) 324–335
4. Schechner, Y.Y., Kiryati, N., Basri, R.: Separation of transparent layers using focus. International Journal of Computer Vision **39** (2000) 25–39
5. Schechner, Y.Y., Kiryati, N., Shamir, J.: Blind recovery of transparent and semireflected scenes. In: Computer Vision and Pattern Recognition. Volume 1. (2000) 38–43

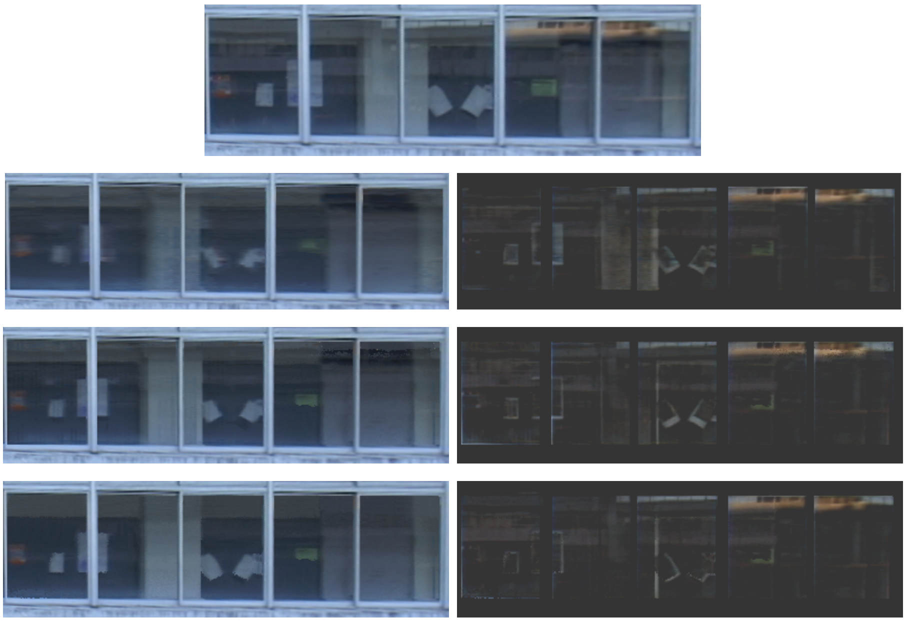


Fig. 9. Input outdoor image and separation results: first row shows the input outdoor image, second row shows the separation results using all EPI range, the results describe in third row are created by using reduce EPI range. Bottom row shows the results by using motion tracking data in EPI rectification process. The right column of the result images are enhanced to make them easier to see.

6. Schechner, Y.Y., Kiryati, N., Shamir, J.: Separation of transparent layers by polarization analysis. In: Scandinavian Conference on Image Analysis. Volume 1. (1999) 235–242
7. Schechner, Y.Y., Shamir, J., Kiryati, N.: Polarization-based decorrelation of transparent layers: The inclination angle of an invisible surface. In: ICCV. Volume 2. (1999) 814–819
8. Schechner, Y.Y., Shamir, J.: Vision through semireflecting media: polarization analysis. Optics Letters **24** (1999) 1088–1090
9. Hermanto, Barros, A.K., Yamamura, T., Ohnishi, N.: Separating virtual and real objects using independent component analysis. IEICE TRANS **E84-D** (2001)
10. Bronstein, A.M., Bronstein, M.M., Zibulevsky, M., Zeevi, Y.Y.: Blind separation on reflections using sparse ICA. In: 4th International Symposium of Independent Component Analysis and Blind Signal Separation. (2003) 227–232
11. Farid, H., Adelson, E.H.: Separating reflections from images using independent components analysis. Journal of the Optical Society of America **16** (1999) 2136–2145
12. Sarel, B., Irani, M.: Separating transparent layers through layer information exchange. In: ECCV. Volume 4. (2004) 328–341