Separation of Reflection Components by Fourier Decoupling

Wei Zhou , Chandra Kambhamettu
Video/Image Modeling and Synthesis (VIMS) Lab
Department of Computer and Information Sciences
University of Delaware, Newark DE 19716, USA,
wzhou,chandra@cis.udel.edu

Abstract

Reflections exist in many natural images. For example, shiny surfaces attempt to reflect the surrounding scene, thus, resulting in a composite image that contains the mixture of reflected light and transmitted light. When viewed from a moving camera, different components appear to move differently relative to each other. Although multiple motion recovery problem has been previously studied by many researchers, few algorithms have been proposed to recover the component images themselves. In this paper, we present a novel approach to accurately separate the individual image components from a sequence of composite images. The separation of the images is realized in the Fourier domain, where each image component is represented by its frequency elements. A minimization framework exploring the magnitude and phase constraints between each frequency element is devised to find the correct motion vectors and decouple the composite frequency elements into individual components. Experimental results using both synthetic and real images demonstrate the efficiency and robustness of our algorithm.

1 Introduction

Reflection is a very common phenomenon in visual imaging process. For example, any shiny or glass-like surface will create a reflected image of its surrounding objects, thus, resulting in a composite image that contains several superimposed components in the scene. Operating computer vision algorithms upon composite images is a tough problem, since the information of one component is highly disturbed by other components. For example, in stereo algorithms, specularity components of the scene will bring up great difficulties in finding the correspondences between the diffuse components of the image pairs. Many attempts have recently been made to decompose the composite image into a set of individual components in order to allow further processing. In [4] [5], the polarization properties of the reflections are utilized, and a linear polarizer is added to the imaging system to help the separation. [6] [7] [8] explore the focus cue of the imaging system, and color information is used in [9] [10] [11] to separate the specularity intensities from the Lambertian intensities.

Since each of the components are originated from different parts of the scene, when viewed from a moving camera, they appear to move differently relative to each other. This brings up the multiple motion recovery problem that has been extensively studied over the past years such as [12][13][14][15]. Although work well in the motion recovery, neither of these studies provide a complete algorithm to recover the component images themselves. In [16], szeliski *etc.* developed a min/max alternation algorithm to recover the component images and the associated motions from composite image sequences. There are several aspects that restrict the robustness of the algorithm. First, the dominant motion must be reliably estimated from the composite image sequences which is very difficult when the crosstalk between each component is significant. The second, the algorithm is very sensitive to image noise and need integer motion shift. In [1], Vernon solved the problem in Fourier domain, and separated the composite images by decoupling the corresponding frequency elements. A closed-form solution is derived in [1] to achieve the separation where each layer is translating with a distinct and unique velocity.

In this paper, we tackle the problem in frequency domain. Instead of just using the magnitude constraint like [1], the phase relationships between each frequence elements are also explored as phase constraint. It ties each individual frequence element together as a global constraint which makes the algorithm much more robust in the presence of image noise. The whole framework is developed in two steps. First, a simple circle fitting algorithm based on magnitude constraint is proposed to separate the component images given known motion estimates, then an error minimization framework exploring both magnitude and phase constraints is devised to recover both component images and motions.

The paper is organized as follows. Section 2 presents the general problem formulation. Section 3 presents the cir-

cle fitting algorithm which separates the component images with known motion estimates. Section 4 presents the error minimization framework for recovering both motion and component images. Sections 5 shows the experiment results and section 6 discusses the conclusion and future work.

2 Composite Image Modeling

In this paper, we assume that each component image is defined by a 2d signal $I_l(x, y)$, which is warped to the current image frame via a motion warping operator $M_l^{(k)}$. Let $M_l^{(k)} \circ f_l$ denotes the warped image. Similar to [1][16], the pure additive model is used to mix the component images:

$$I^{(k)} = \sum_{l=0}^{L-1} M_l^{(k)} \circ I_l \tag{1}$$

From equation 1, in order to extract the component images I_l from composite image sequence $I^{(k)}$, the motions $M_l^{(k)}$ should be reliably estimated, which is a tough issue itself in the computer vision field. In spite of putting the burden on the multiple motion estimator, we simplified the problem by restricting the camera only to translate along a plane parallel to the camera image plane when capturing the image sequences. In practice, this is not a hard operation. Under this setting, the motions of each component images can be simply approximated by the pure translations. The proof is simple. Let $z1$, $z2$ be the depths of two points in the scene from the camera and $d1$, $d2$ be the associated motion vectors (It is also known as disparity in the stereo vision). Since disparity is proportional to the inverse of the depth, that is $d1 \propto 1/z1$ and $d2 \propto 1/z2$, subtracting the two equations from each other, we get $\Delta d \propto (\Delta z/z1z2)$. If $\Delta z \ll z1z2$, $\Delta d \to 0$. So, when the depth range of the object is small enough compared to the distance from the camera, the motion difference between points of the object could be ignored, and the motion of the whole object can be approximated as pure translation. Based on this, the equation 1 can be re-written as

$$I^{(k)}(x, y) = \sum_{l=0}^{L-1} I_l(x - TX_l^{(k)}, y - TY_l^{(k)}) \tag{2}$$

where $(TX_l^{(k)}, TY_l^{(k)})$ denotes the translation vectors. Transform it to the frequency domain, we get:

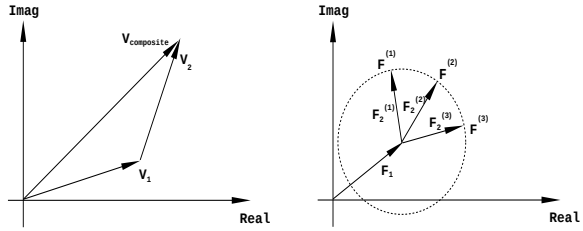
$$F^{(k)}(u, v) = \sum_{l=0}^{L-1} F_l(u, v) e^{-j(\frac{2\pi TX_l^{(k)} u}{M} + \frac{2\pi TY_l^{(k)} v}{N})} \tag{3}$$

So, the separation can be done in the frequency domain by decoupling the composite Fourier components F^k into several individual Fourier components F_l. Then the required component images can be obtained by the inverse Fourier transform. In the following sections, a novel algorithm is proposed to solve the problem. To make the idea more clear, from now on, we will focus on the situation where only 2 component images exist in the composite images. As we will discuss later, the idea can be easily extended to handle arbitrary number components separation.

3 Circle Fitting Algorithm

An intuitive way to analyze signal in Fourier domain is to view each frequency element as a vector in the complex plane. Each frequency element of the composite image is just a composite vector obtained by adding up two component vectors, as shown in Figure 1a.

$$V_{composite} = V_1 + V_2 \tag{4}$$



(a) (b)

Figure 1: Fourier analysis in complex plane

Let's first consider the case where only one component image is moving. Figure 1b shows the case when three composite images are available and only component 2 is translating. According to the translation property of the Fourier transform, the translation in time domain is equivalent to phase change in frequency domain; the magnitudes of the frequency elements will not change, so:

$$\|F_2^{(1)}(u, v)\| = \|F_2^{(2)}(u, v)\| = \|F_2^{(3)}(u, v)\|. \tag{5}$$

If we fit a circle using $F^{(1)}$, $F^{(2)}$, and $F^{(3)}$, the origin of the circle should be F_1. Now, it is clear that the image separation problem can be easily solved in frequency domain by a simple circle fitting algorithm. We know that in order to fit a circle, at least 3 points must be provided. So, at least 3 composite images should be taken for the successful separation.

So far so good, but in close inspection, we find the algorithm has two main limitations:

1. Noise sensitivity. The circle fitting algorithm is very sensitive to noise. In figure 2, we show that if we fit a circle

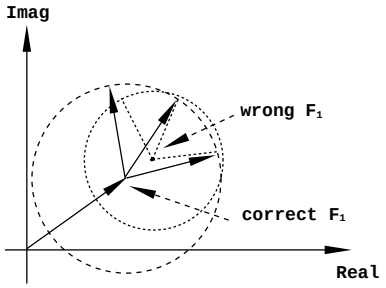


Figure 2: Circle fitting algorithm is noise-sensitive

using the noisy points, the resultant origin could be very far away from the correct one.

2. Phase periodic problem. The relationship between phase change $\Delta\phi$ in frequency domain and translation vector (TX, TY) in time domain is:

$$\Delta\phi(u,v) = 2\pi(\frac{TXu}{M} + \frac{TYv}{N}) \tag{6}$$

For some (TX, TY) and (u, v) combinations, we may have $(\frac{TXu}{M} + \frac{TYv}{N}) = k, k \in N$, so $\Delta\phi(u,v) = 2k\pi$. In this case, the vector rotates several rounds and stops at the original position. We call this (u, v) the periodic frequency. Apparently, we can't have three distinct vectors in the complex plane at periodic frequency. Only two or may be one (the worst case) distinct vectors are available, in which case, the circle fitting algorithm will definitely fail.

Figure 3 demonstrates the separation results using the circle fitting algorithm under different situations. We can see that as noise or number of periodic frequency increase, the results are getting worse.

The reason that the circle fitting algorithm tends to fail is because it only uses the magnitude constraint. The information inside the phase relationship between different frequencies is ignored. In the next section we will explore the phase constraint in detail, and based on that, a more sophisticated algorithm is proposed. Experiment results show that it works better and more robust than the simple circle fitting algorithm.

4 Error Minimization Framework

4.1 Incorporating the phase constraint

The phase changes of different frequency elements (u, v) are not unrelated. Actually, they are related by the means of translation vectors. By equation 6, if we know the translation vector (TX, TY), the phase changes of all frequency elements must comply with equation 6, we call it the phase constraint.

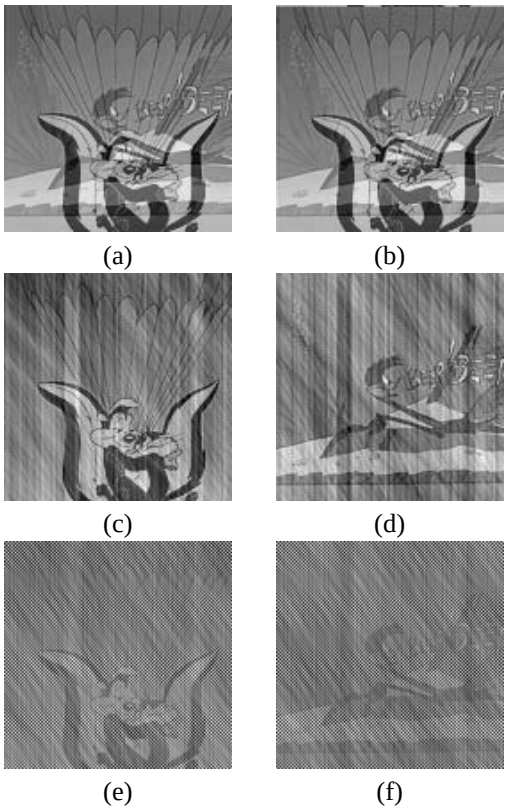


Figure 3: Image separation results using circle fitting algorithm under different situations. (a) and (b) are two of the three layered images, (c) and (d) show the case when there are 4 periodic frequencies, (e) and (f) are the results when 4 periodic frequencies are present and additional Gaussian noise is added ($\mu = 0$, $\sigma = 5$).

For sake of brevity, we drop the (u, v) from now on, but remember that the equations hold for all frequency elements. Using the notations of figure 4, by equation 3, we have:

$$F_2^{(2)} = F_2^{(1)} e^{-j\Delta\phi_2^{(2)}} \tag{7}$$

$$F_2^{(3)} = F_2^{(1)} e^{-j\Delta\phi_2^{(3)}} \tag{8}$$

replace $F_2^{(k)}$ with $F^{(k)} - F_1$, we get:

$$F^{(2)} - F_1 = (F^{(1)} - F_1) e^{-j\Delta\phi_2^{(2)}} \tag{9}$$

$$F^{(3)} - F_1 = (F^{(1)} - F_1) e^{-j\Delta\phi_2^{(3)}} \tag{10}$$

rearrange the terms of above two equations:

$$F_1(1 - e^{-j\Delta\phi_2^{(2)}}) = F^{(2)} - F^{(1)} e^{-j\Delta\phi_2^{(2)}} \tag{11}$$

$$F_1(1 - e^{-j\Delta\phi_2^{(3)}}) = F^{(3)} - F^{(1)} e^{-j\Delta\phi_2^{(3)}} \tag{12}$$

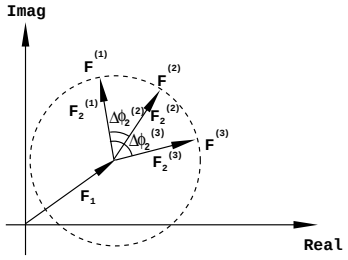


Figure 4: Phase change constraint

Suppose the translation vectors $(TX_2^{(2)}, TY_2^{(2)})$ and $(TX_2^{(3)}, TY_2^{(3)})$ are known, the phase changes $\Delta\phi_2^{(k)}$ can be calculated by equation 6, then above two equations only have one unknown F_1. Obviously, this is an overconstraint problem with the following matrix form:

$$\begin{bmatrix} 1 - e^{-j\Delta\phi_2^{(2)}} \\ 1 - e^{-j\Delta\phi_2^{(3)}} \end{bmatrix} F_1 = \begin{bmatrix} F^{(2)} - F^{(1)} e^{-j\Delta\phi_2^{(2)}} \\ F^{(3)} - F^{(1)} e^{-j\Delta\phi_2^{(3)}} \end{bmatrix} \quad (13)$$

The best F_1 can be obtained by least mean square (LMS) solution.

4.2 Both component images moving simultaneously

When two images are translating simultaneously, we have 4 unknown translation vectors instead of 2. Figure 5 shows the relationship between the vectors in the complex plane.

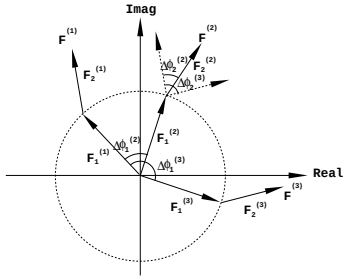


Figure 5: Normal case

Although the relationships between the vectors shown in figure 5 are not so clear as in figure 4, they still satisfy the two constraints. i.e., the amplitude of F_1 and F_2 will not change, and also the phase changes of F_1 and F_2 will comply with the phase constraint. Combining both constraints, we get similar equations as 9, 10:

$$F^{(2)} - F_1^{(1)} e^{-j\Delta\phi_1^2} = (F^{(1)} - F_1^{(1)}) e^{-j\Delta\phi_2^2} \quad (14)$$

$$F^{(3)} - F_1^{(1)} e^{-j\Delta\phi_1^3} = (F^{(1)} - F_1^{(1)}) e^{-j\Delta\phi_2^3} \quad (15)$$

Assuming known translation vectors, $F_1^{(1)}$ could be calculated by LMS.

4.3 Final implementation

If the given translation vectors are correct, the LMS algorithm should be better than the circle fitting algorithm because it uses both constraints. So the crucial part of the algorithm becomes the estimation of the translation vectors. Notice that the LMS method is trying to best satisfy both constraints. If the given translation vectors are wrong, we would get the wrong phase constraint, and the answer would be pulled from the correct position to some wrong position in the solution space. In consideration of this, we define two error functions to measure the deviations of the results from the two constraints:

The error function for magnitude constraint is straightforward:

$$E_a = \sum_l \sum_{k1 \neq k2} std_{(u,v)}(\|F_l^{(k1)}(u,v)\| - \|F_l^{(k2)}(u,v)\|) \quad (16)$$

where $std_{(u,v)}$ means standard deviation operation over all (u, v). So, the magnitude errors of all frequencies (u, v) are calculated first, then the standard deviation of the errors is calculated and summed over all frames and components.

In order to measure the deviation of the results away from the phase constraint, we first fit the current phase change estimates using the phase constraint equations (6), and then calculate the deviation of the current estimates from the fitted model. Suppose the current estimates of the phase changes of F_l are $\Delta\phi_l^{(k)}$. Fitting it into the phase constraint, we get the fitted phase model $\Delta\phi_l^{(k)'}$. The error function for the phase constraint is defined as:

$$E_p = \sum_l \sum_k std_{(u,v)}(\Delta\phi_l^{(k)}(u,v) - \Delta\phi_l^{(k)'}(u,v)) \quad (17)$$

The total error function is defined as:

$$E = \lambda_1 E_a + \lambda_2 E_p \quad (18)$$

where λ_1 and λ_2 are two weights to balance the two constraints. In the experiments, we choose $\lambda_1 = 0.85$ and $\lambda_2 = 0.15$, and the results are satisfactory .

Obviously, the larger value of the error function accounts for the larger error in the translation vector estimation, and vise-versa. Based on this, an error minimization framework is proposed to get the correct estimation of the translating

vectors, and at the same time the best separation of the images. The whole algorithm is outlined in the following four steps:

1. By 2D FFT, transform the three layered images into the frequency domain.

2. In case of only one component moving, the frequency elements is initially decoupled by circle fitting algorithm, and then initial values of the translation vectors is calculated from it. In the other case, we initialize the translation vectors randomly.

3. LMS method is applied to decouple the frequency elements using the current values of the translation vectors, and the result is used to calculate the error function.

4. Downhill simplex minimization method is used to minimize the error. The final decoupled frequency elements with minimal error are transformed back to the time domain using inverse Fourier transform.

5 Experiment results

Applicability of the proposed algorithm has been tested using both synthetic and real images. Figure 6 shows the experimental results using synthetic images. The composite images are formed by adding two cartoon images which translate over time. Figures 6c and 6d show the results when 4 periodic frequencies exist. Figures 6e and 6f are the separation results when additional Gaussian noise is added ($\mu = 0, \sigma = 5$). The results are much better and more robust compared to the circle fitting results (figure 3c,d,e,f).

Experiment results on real images are shown in figure 7 and 8. In Figure 7, a glass with pattern "H" is placed on top of a book page, 3 images are captured using canon powershot s30 camera when the glass is translating around; as the result shows, the book page and the glass pattern are successfully separated. Figure 8 shows the case when two components are moving simultaneously; we photographed a picture framed behind glass with the reflection of a water bottle. The picture with reflection was photographed three times with camera translating around. As we see in figure 8a and 8b, both the picture behind glass and the reflection of water bottle are translating when the camera is translating. Figure8c and 8d show the separation results. Due to the curvedness of the glass surface and the bottle shape, Some deformation occurs during the translating. But the results are still promising.

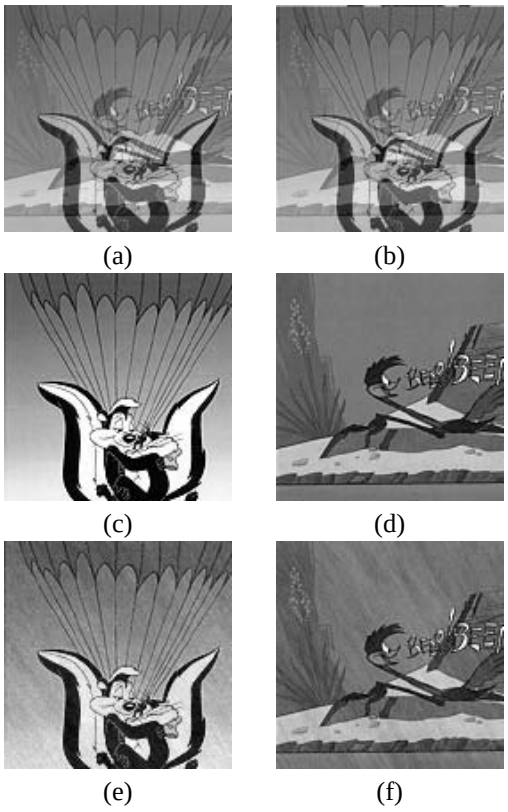


Figure 6: Experiments on synthetic images. The layered images are formed by adding two cartoon images which translate over time. (a)and (b) show two of three layered images. (c) and (d) show the separation result when 4 periodic frequencies exist. (e) and (f) are the separation results when additional Gaussian noise is added ($\mu = 0, \sigma = 5$).

6 Conclusion and future work

This paper presents a novel scheme to separate a sequence of composite images into individual component images by Fourier components decoupling. The main contribution of this work is that both the magnitude and phase constraints between different frequency elements are explored, and an error minimization framework combining these two constraints is proposed to recover the component images and the motion. Experiment results demonstrate the efficiency and robustness of the algorithm.

The work can be extended in several directions. As we mentioned before, although we focused on the two component images separation, the algorithm can be easily extended to arbitrary number component image separation simply by adding more translation unknows into the error functions; the magnitude and phase constraints are always valid. More complex motion other than translations could be solved by dividing the images into small areas and assuming a simple translation within the small area. All of these will be considered in our future work.

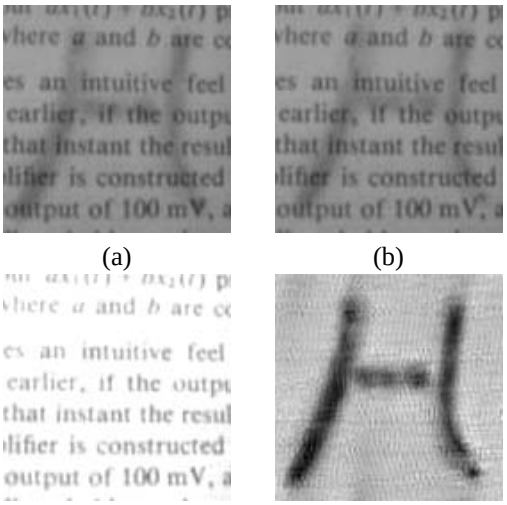


Figure 7: Experiments on real images. A glass with pattern "H" is placed on top of a book page, (a) and (b) show two of three captured layered images with glass translating around, (c) and (d) are the separation results.

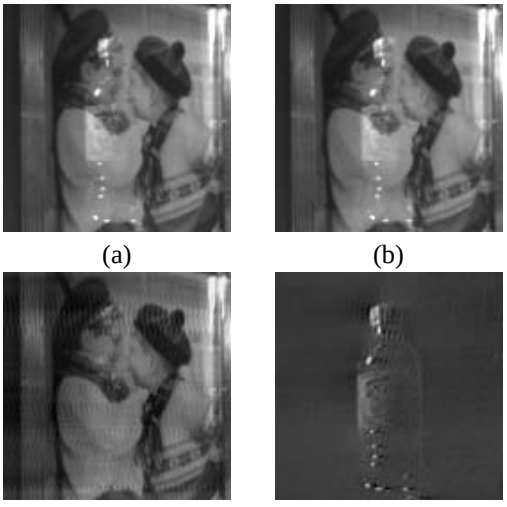


Figure 8: Two layers translating simultaneously. Three images of a picture plus the reflection of a water bottle is captured with a moving camera, (a) and (b) show two of three layered images, (c) and (d) are the separation results

References

[1] David Vernon. Decoupling Fourier Components of Dynamic Image Sequences: A Theory of Signal Separation, Image Segmentation, and optical Flow Estimation. *Proceedings of the 5th European Conference on Computer Vision - ECCV'98*, volume 2, pp. 69-85.

[2] David Vernon. Segmentation in dynamic image sequences by isolation of coherent wave profiles. *Proc. ECCV'96* volume 1, pp. 293-303.

[3] David Vernon Fourier Vision–Segmentation and Velocity Measurement Using the Fourier Transform. *Kluwer Academic Publishers*

[4] H.Farid, E.H.Adelson. Separating Reflections and Lighting Using Independent Components Analysis. *IEEE CVPR*, (1999) volume 1, pp.262-267

[5] N.Ohnishi, K.Kumaki, T.Yamanura, T.Tanaka. Separating real and virtual objects from their overlapping images. *Proc. ECCV'96* volume 2, pp636-646

[6] Y.Y.Schechner, N.Kiryati, J.Shamir. Blind Recovery of Transparent and Semireflected Scenes. *IEEE CVPR*, (2000) volume 1, pp.38-43

[7] Y.Y.Schechner, N.kiryati, R.Basri. Separation of transparent layers using focus. *Proc. ICCV*, (1998) pp. 1061-1066

[8] Y.Y.Schechner, J.Shamir, N.Kiryati. Polarization-based decorrelation of transparent layers: The inclination angle of an invisible surface. *Proc.ICCV*, (1999) volume 2, pp. 814-819.

[9] S.Shafer. Using color to separate reflection components. *Color research and applications*, 10:210-218, 1985.

[10] S.Nayar, X.Fang, B.Terrance. Separation of reflection components using color and polarization. *International Journal of Computer Vision*, (1997) 21(3):163-186.

[11] S.Lin, H-Y.Shum. Separation of Diffuse and Specular Reflection in Color Images. *IEEE CVPR* (2001) volume 1, PP.341-346

[12] J.Y.A.Wang, E.H.Adelson. Layered representation from motion analysis. *IEEE CVPR* (1993) PP.361-365

[13] Y.Weiss. Smoothness in layers: Motion segmentation using nonparametric mixture estimation. *IEEE CVPR* (1997) PP.520-526

[14] S.Baker, R.Szeliski, P.Anandan. A layered approach to stereo reconstruction. *IEEE CVPR* (1998), PP.434-441.

[15] J.R.Bergen, P.J.Burt, R.Hingorani, S.Peleg. A three-frame algorithm for estimating two-component image motion. *IEEE Trans. PAMI* 14, (1992) PP.886-896.

[16] R.Szeliski, S.Avidan, P.Anandan Layer extraction from multiple images containing reflections and transparency *IEEE CVPR* (2000), PP.246-253.