Volume 2, Issue 4, April 2012 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

An Image Alignment Based on Enhanced Correlation Coefficient

Deepika Dubey, Abhishek Jain, Uday Pratap Singh
LNCT-Bhopal, India
sweetdeepika1087@gmail.com

Abstract:-In this work we present an overview of image alignment, describing most of the algorithms and their extensions in a consistent framework. We concentrate on the inverse compositional algorithm, an efficient algorithm that we recently proposed. We cover the quantity approximated, the warp update rule, and the gradient descent approximation. In this work, we propose the use of a modified version of the correlation coefficient as a performance criterion for the image alignment problem. The proposed modification has the desirable characteristic of being invariant with respect to photometric distortions. Since the resulting similarity measure is a nonlinear function of the warp parameters, we develop two iterative schemes for its maximization, one based on the forward additive approach and the second on the inverse compositional method. As is customary in iterative optimization, in each iteration, the nonlinear objective function is approximated by an alternative expression for which the corresponding optimization is simple. In our case, we propose an efficient approximation that leads to a closed-form solution (per iteration) which is of low computational complexity, the latter property being particularly strong in our inverse version. The proposed schemes are tested against the Forward Additive Lucas-Kanade and the Simultaneous Inverse Compositional (SIC) algorithm through simulations. Under noisy conditions and photometric distortions, our forward version achieves more accurate alignments and exhibits faster convergence, whereas our inverse version has similar performance as the SIC algorithm but at a lower computational complexity.

Keywords: Image alignment, Lucas-Kanade, a unifying framework, additive vs. compositional algorithms, for- wards vs. inverse algorithms, the inverse compositional algorithm, efficiency, steepest descent, Gauss-Newton, Newton, Levenberg-Marquardt.

1. INTRODUCTION

The alignment problem can be seen as a mapping between the coordinate systems of two images; therefore, the first step toward its solution is the suitable selection of a geometric transformation that adequately models this mapping. Existing models are basically parametric [12] and their exact form heavily depends on the specific application and the strategy selected to solve the alignment problem [3], [13]. The class of affine transformations and, in particular, several special cases (as pure translation) have been the center of attention in many application, [4,5,10]. Alternative approaches rely on projective transformations (homography) and, more generally, on nonlinear transformations [5] [15].

A common assumption encountered in most existing techniquesis the brightness constancy of corresponding points or regions in the two profiles [20]. However, this assumption is valid only in specific cases and it is obviously violated under varying illumination conditions. There, it becomes clear that, in a practical situation, it is important that the alignment algorithm be able to take into account illumination changes. Alignment techniques that compensate for photometric distortions in contrast and brightness have been proposed in [1], [6], [8], [10], [16]. Alternative schemes make use of a set of basis images for handling arbitrary lighting conditions [3], [21] or use spatially dependent photometric models [7].

Image alignment consists of moving, and possibly deforming, a template to minimize the difference between the template and an image. Since the first use of image alignment in the Lucas- Kanade optical flow algorithm [13], image alignment has become one of the most widely used techniques in computer vision. Besides optical flow, some of its other applications include tracking [5, 12], parametric and layered motion estimation [4], mosaic construction [16], medical image registration [7], and face coding [2, 8].

The usual approach to image alignment is gradient descent. A variety of other numerical algorithms such as *difference decomposition* [11] and *linear regression* [8] have also been proposed, but gradient descent is the defacto standard. Gradient descent can be performed in variety of different ways, however. One difference between the various approaches is whether they estimate an additive increment to the parameters (the *additive* approach [13]), or whether they estimate an incremental warp that is then composed with the current estimate of the warp (the *compositional* approach [16].) Another difference is whether the algorithm performs a Gauss-Newton, a Newton, a steepest-descent, or a Levenberg-Marquardt approximation in each gradient descent step.

We propose a unifying framework for image alignment, describing the various algorithms and their extensions in a consistent manner. Throughout the framework we concentrate on the *inverse compositional* algorithm, an efficient algorithm that we recently proposed [2]. We examine which of the extensions to Lucas-Kanade can be applied to the inverse compositional algorithm without any significant loss of efficiency, and which extensions require additional computation. Wherever possible we provide empirical results to illustrate the various algorithms and their extensions.

We categorize algorithms as either *additive* or *compositional*, and as either *forwards* or *inverse*. We prove the first order equivalence of the various alternatives, derive the efficiency of the resulting algorithms, describe the set of warps that each alternative can be applied to, and finally empirically compare the algorithms. In Section 4 we describe the various gradient descent approximations that can be used in each iteration, *Gauss-Newton*, *Newton*, *diagonal Hessian*, *Levenberg-Marquardt*, and *steepest-descent* [14]. We compare these alternatives both in terms of speed and in terms of empirical performance. We conclude in Section 5 with a discussion. In future papers in this series (currently under preparation), we will cover the choice of the error norm, how to allow linear appearance variation, how to add priors on the parameters, and various techniques to avoid local minima.

A common assumption encountered in most existing techniques is the brightness constancy of corresponding points or regions in the two profiles [20]. However, this assumption is valid only in specific cases and it is obviously violated under varying illumination conditions. There, it becomes clear that, in a practical situation, it is important that the alignment algorithm be able to take into account illumination changes. Alignment techniques that compensate for photometric distortions in contrast and brightness have been proposed in [1], [6], [8], [10], [16]. Alternative schemes make use of a set of basis images for handling arbitrary lighting conditions[3], [21] or use spatially dependent photometric models [7].

we adopt a recently proposed similarity measure [11], the enhanced correlation coefficient, as our objective function for the alignment problem. Our measure is characterized by two very desirable properties. First, it is invariant to photometric distortions in contrast and brightness. Second, although it is a nonlinear function of the parameters, the iterative scheme we are going to develop for the optimization problem will turn out to be linear, thus requiring reduced computational complexity. Despite the resemblance of our final algorithm to well-known variants of the Lucas-Kanade alignment method which take lighting changes into account [10], [19], its performance, as we are going to see, is notably superior. We would like to mention that the enhanced correlation coefficient criterion was successfully applied to the problem of 1D translation estimation in stereo correspondence [11] and 2D translation estimation in registration [2].

1.2 History

Algorithms for aligning images and stitching them into seamless photo-mosaics are among the oldest and most widely used in computer vision. Frame-rate image alignment is used in every camcorder that has an “image stabilization” feature. Image stitching algorithms create the high- resolution photo-mosaics used to produce today’s digital maps and satellite photos. They also come bundled with most digital cameras currently being sold, and can be used to create beautiful ultra wide-angle panoramas.

An early example of a widely-used image registration algorithm is the patch-based translational alignment (optical flow) technique developed by Lucas and Canada (1981). Variants of this algorithm are used in almost all motion-compensated video compression schemes such as MPEG andH.263 (Le Gall 1991). Similar parametric motion estimation algorithms have found a wide variety of applications, including video summarization (Bergen et al. 1992a, Theodosia and Bender 1993,Kumar et al. 1995, Iran and Anandan 1998), video stabilization (Hansen et al. 1994), and video compression (Irani et al. 1995, Lee et al. 1997). More sophisticated image registration algorithms have also been developed for medical imaging and remote sensing—see (Brown 1992, Zitov’aa and Flusser 2003, Goshtasby 2005) for some previous surveys of image registration techniques.

In the photo grammetry community, more manually intensive methods based on surveyed ground control points or manually registered tie points have long been used to register aerial photos into large-scale photo-mosaics (Slama 1980). One of the key advances in this community was the development of bundle adjustment algorithms that could simultaneously solve for the locations of all of the camera positions, thus yielding globally consistent solutions (Triggs et al. 1999). One of the recurring problems in creating photo-mosaics is the elimination of visible seams, for which a variety of techniques have been developed over the years (Milgram 1975, Milgram 1977, Peleg 1981, Davis 1998, Agarwala et al. 2004).

1.3 Application

Image alignment is the process of matching one image called template (let's denote it as T) with another image. There are many applications for image alignment, tracking object on video , motion estimation analysis, and other tasks of computer vision , such as object tracking , image registration problem , region tracking.

1.3.1 Image registration:-

Image registration is the process of aligning two or more images of the same scene. Typically, one image, called the *base* image or *reference* image, is considered the reference to which the other images, called *input* images, are compared. The object of image registration is to bring the input image into alignment with the base image by applying a spatial transformation to the input image. The differences between the input image and the output image might have occurred as a result of terrain relief and other changes in perspective when imaging the same scene from different viewpoints. Lens and other internal sensor distortions, or differences between sensors and sensor types, can also cause distortion.

1.3.2 Object tracking:-

Object tracking consists in estimation of trajectory of moving objects in the sequence of images. Automation of the computer object tracking is a difficult task. Dynamics of multiple parameters changes representing features and motion of the objects, and temporary partial or full occlusion of the tracked objects have to be considered. This monograph presents the development of object tracking algorithms, methods and systems. Both, state of the art of object tracking methods and also the new trends in research are described in this book. Fourteen chapters are split into two sections. Section 1 presents new theoretical ideas whereas Section 2 presents real-life applications. Despite the variety of topics contained in this monograph it constitutes a consisted knowledge in the field of computer object tracking. The intention of editor was to follow up the very quick progress in the developing of methods as well as extension of the application.

1.3.3 Motion estimation:-

Successive video frames may contain the same objects (still or moving). Motion estimation examines the movement of objects in an image sequence to try to obtain vectors representing the estimated motion. Motion compensation uses the knowledge of object motion so obtained to achieve data compression. In interface coding, motion estimation and compensation have become powerful techniques to eliminate the temporal redundancy due to high correlation between consecutive frames.

1.3.4 Computer vision:-

Computer vision is a field that includes methods for acquiring, processing, analyzing, and understanding images and, in general, high-dimensional data from the real world in order to produce numerical or symbolic information, e.g., in the forms of decisions. A theme in the development of this field has been to duplicate the abilities of human vision by electronically perceiving and understanding an image. This image understanding can be seen as the disentangling of symbolic information from image data using models constructed with the aid of geometry, physics, statistics, and learning theory.

2. Outline

In this section we briefly introduce the problem of alignment of two image profiles. To this end, let us assume that a *reference* image I_r (x) and a *warped* image I_W ($\mathrm{x}^{\emptyset}$) are given, where x = [x, y] and $\mathrm{x}^{\emptyset}$ = [$x^{\emptyset}$, $y^{\emptyset}$] denote coordinates. Suppose also that we are given a set of coordinates $S = \{\mathrm{x}_i \mid i = 1, \ldots, K\}$ in the reference image, which is called *target area*. Then, the alignment problem consists in finding the corresponding coordinate set in the warped image. By considering that a transformation model T (x; p) where

$$\mathrm{p} = (p_1, p_2, \ldots, p_N)^t$$

is a vector of unknown parameters is given, the alignment problem is reduced to the problem of estimating the parameter vector p such that

$$I_r(\mathrm{x}) = \Psi(I_W(T(\mathrm{x}; \mathrm{p})); \alpha), \quad \mathrm{x} \in S, \qquad (1)$$

where transformation $\Psi(I\ ,\ \alpha)$ which is parameterized by a vector α, accounts for possible photometric distortions that violate the brightness constancy assumption, a case which arises in real applications due to different viewing directions and/or different illumination conditions.

The goal of most existing algorithms is the minimization of the dissimilarity of the two image profiles, providing the optimum parameter values. Dissimilarity is usually expressed through an objective function $E(\mathrm{p}, \alpha)$ which involves the l_p norm of the intensity residual of the image profiles. A typical minimization problem has the following form

$$\min E(\mathrm{p}, \alpha) = \min \sum |I_r(\mathrm{x}) - \Psi(I_W(T(\mathrm{x}; \mathrm{p})), \alpha)|^p$$

Solving the above defined problem is not a simple task because of the nonlinearity involved in the correspondence part. Computational complexity and estimation quality of existing schemes depends on the specific l_p norm and the models used for warping and photometric distortion.

3. Problem formation

In this section we briefly introduce the problem of alignment of two image profiles. To this end, let us assume that a *reference* image I_r (x) and a *warped* image I_W (x) are given, where x = [x, y] and x = [x , y] denote coordinates. Suppose also that we are given a set of coordinates $S = \{\mathrm{x}_i \mid i = 1, \ldots, K\}$ in the reference image, which is called *target area*. Then,

the alignment problem consists in finding the corresponding coordinate set in the warped image.
By considering that a transformation model $T(\mathrm{x}; \mathrm{p})$ where $\mathrm{p} = (p_1, p_2, \ldots, p_N)^t$ is a vector of unknown parameters is given, the alignment problem is reduced to the problem of estimating the parameter vector p such that

$$I_r(\mathrm{x}) = \Psi(I_w(T(\mathrm{x}; \mathrm{p})); \alpha), \ \mathrm{x} \in S,$$

where transformation $\Psi(I, \alpha)$ which is parameterized by a vector α, accounts for possible photometric distortions that violate the brightness constancy assumption, a case which arises in real applications due to different viewing directions and/or different illumination conditions.
The goal of most existing algorithms is the minimization of the dissimilarity of the two image profiles, providing the optimum parameter values. Dissimilarity is usually expressed through an objective function $E(\mathrm{p}, \alpha)$ which involves the l_p norm of the intensity residual of the image profiles. A typical minimization problem has the following form Solving the above defined problem is not a simple task because of the nonlinearity involved in the correspondence part. Computational complexity and estimation quality of existing schemes depends on the specific l_p norm and the models used for warping and photometric distortion. As far as the norm power p is concerned most methods use $p = 2$ (Euclidean norm). This will also be the case in the approach we briefly present in the next section.

4. PROPOSED CRITERION AND MAIN RESULTS

Under the warping transformation ð(x; p),the quardinate xk, k=1………,k and denote with i_r and i_w ðpÞ their zero-mean versions, which are obtained by subtracting from each vector its corresponding arithmetic mean. We then propose the following criterion to quantify the performance of the warping transformation with parameters p.

4.1 Performance Measure Optimization

Once the performance measure is specified, we then continue with its minimization in order to compute the optimum parameter values. It is straightforward to prove that minimizing E_{ECC} ðpÞ is equivalent to maximizing the version of the zero-mean reference vector, which is constant. Notice that, even if i_w ðpÞ depends linearly on the parameter vector p, the resulting objective function is still nonlinear with respect to p due to the normalization of the warped vector. This, of course, suggests that its maximization requires nonlinear optimization techniques
Performed either by using direct search or by gradient-based approaches. Here, we are going to use the latter. As is customary in iterative techniques, we are going to replace the original optimization problem with a sequence of secondary optimizations. Each secondary optimization relies on the outcome

5. Simulation Result

In this section, we perform a number of simulations in order to evaluate our FA-ECC and IC-ECC algorithmic version. As we mentioned above, we will also simulate the FA-LK algorithmic version that copes with photometric distortions and the SIC algorithm, which is considered to be the most effective inverse LK scheme. For all aspects affecting the simulation experiments, we made an effort to stay exactly within the framework specified in [13], [19]. To model the warping process, we are going to use the class of affine transformations. We know that the 2D rigid body or similarity transformation are members of this class. Furthermore, the Jacobean of the affine model is a constant matrix, meaning that it can be computed offline. Before proceeding with the presentation of our simulation results, let us first briefly present the experimental setup and the figures of merit we are going to adopt.

5.1 First Experiment

In this experiment, for the intensity noise, we use a standard deviation i, which corresponds to eight gray levels, and compare the convergence characteristics of the competing algorithms for a maximum number of iterations^2 jmax ¼ 15 and TMSD ¼ 1 pixel^2. Figs. 1a, 1b, and 1c depict the convergence profiles of the algorithms for different values of p. We observe the appearance of an MSD floor value in each algorithm which is due to the presence of the intensity noise. Fig. 1d presents the corresponding PoC as a function of p .As we can see, each algorithm attains a different MSD floor value with our FA-ECC version converging to the lowest one and with a rate which can be significantly better. Specifically, for weak geometric deformations, all algorithms reach almost comparable floor values and have comparable convergence rates, with FA-ECC being slightly faster than its rivals. However, in the case of medium to strong deformations, FA-ECC reaches an MSD floor value which is 3 dB lower than the inverse versions and slightly lower than the FA-LK algorithm. On the other hand, the convergency rate of FA-ECC is significantly superior compared to all other algorithms. Regarding our IC-ECC version, as we can see, it has performance comparable to the SIC algorithm. The same characteristics also apply to PoC, where FA-ECC exhibits a larger percentage of successful convergences while IC-ECC matches the performance of SIC. Regarding the third figure of merit, we applied the algorithms for a

maximal number of iterations $j_{MAX} = 100$. In order to test the accuracy of the alignment, we selected a threshold value $T_{MSD} = (1/18 \text{ pixel})^2$ (i.e., 25 dB), assuring that T_{MSD} is higher than the MSD floor value of all competing algorithms. Fig. 3a depicts the corresponding curves for three values of p. As we can see, for weak deformations, all algorithms are almost com pletely successful after the 10th iteration. When, however, the geometric deformation be-comes stronger, FA-ECC outperforms its competitors significantly. Again, IC-ECC is comparable to SIC.

5.2 Second Experiment

In this simulation, we consider the realistic case of photometrically distorted images under noisy conditions. We consider two different scenarios. We impose the photometric distortion 1) on the reference image and 2) on the warped one. Since all competing algorithms perfectly compensate for linear photometric distortions, we consider a nonlinear transformation of the for We repeat the same set of simulations as in the first experiment, only now we impose the photometric distortion before adding intensity noise.

The results we obtained are shown in Fig. 2. As we can see, the performance of our forward algorithm seems to be almost unaffected, achieving, under both scenarios, almost the same and the lowest MSD floor value. On the other hand, the performance of both inverse algorithms and FA-LK scheme seems to be vitally affected. Comparing Fig. 2 with Fig. 1, we observe that, under the first scenario, FA-ECC performs even better than before. In fact, the MSD floor value is now 3 and 5 dB lower than the value attained by the FA-LK algorithm and the inverse algorithms, respectively. We should note here that the MSD floor is due not only to the intensity noise but also to the photometric model mismatch. Under the second scenario, all algorithms achieve the same MSD floor value. As far as PoC is concerned, we observe a rather steady and robust behavior for the forward algorithms under both scenarios while inverse schemes, under the first scenario, exhibit a significant performance reduction as compared to the second one. Again, FA-ECC outperforms the other algorithms. Com- paring Fig. 3a with Fig. 3b, we can also notice a robust and consistent behavior of FA-ECC with respect to intensity noise and photometric distortion model mismatch.

In summary, we can safely conclude that our proposed schemes are preferable to the corresponding variants of the LK algorithm. Clearly, our forward version is more effective than the forward LK scheme regarding both speed and percentage of convergence. On the other hand, our inverse version has performance which is comparable to the performance of SIC, which is the best inverse version of the LK algorithm. However, the point that makes our IC-ECC version preferable to SIC is the reduced computational complexity.

We should also mention that we evaluated the algorithms under diverse uncertainty conditions. Only in the case of zero intensity noise (in other words, when the warped image follows the warping model exactly), we observed the performance of both inverse algorithms and the FA-ECC to be similar to that one.

addition, based on the inverse compositional update rule, we developed an efficient modification of the forward algorithm. Our iterative schemes were compared against two variants of the LK algorithm through numerous simulations. Under ideal conditions ,the proposed algorithms and the SIC algorithm exhibited similar performance, out performing the forward LK algorithm. However, in the more realistic case of noisy conditions and photometric distortions, our forward algorithm exhibited a noticeably superior performance in convergence speed, accuracy, and percentage of convergence.

6. Conclusion

In this paper, we have proposed a new I2-based iterative algorithm tailored to the parametric image alignment problem. The new scheme is aimed at maximizing the Enhanced Correlation Coefficient function, which constitutes a measure that is robust against geometric and photometric distortions. The optimal parameters were obtained by iteratively solving a sequence of approximate nonlinear optimization problems which enjoy a simple closed-form solution with low computational cost. In addition, based on the inverse compositional update rule, we developed an efficient modification of the forward algorithm. Our iterative schemes were compared against two variants of the LK algorithm through numerous simulations. Under ideal conditions, the proposed algorithms and the SIC algorithm exhibited similar performance, outperforming the forward LK algorithm. However, in the more realistic case of noisy conditions and photometric distortions, our forward algorithm exhibited a noticeably superior performance in convergence speed, accuracy, and percentage of convergence.

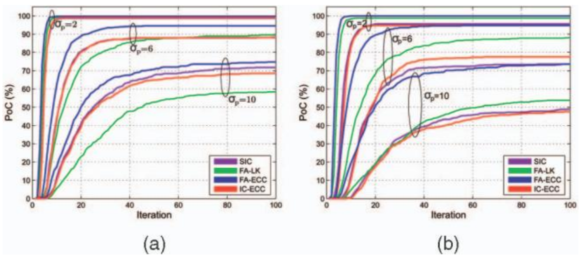

Fig. 1. PoC as a function of iterations: (a) noisy images (b)noisy i

The first inequality is true because of the non positivity of u ɑ^t v (from our assumption); for the second, we applied the Schwartz inequality in the numerator; finally, for the last, we used the fact that the ratio is smaller than 1. We observe that, in this case, we end up with a different (smaller) upper bound. In order to verify its tightness (i.e., whether it constitutes a supremum), we use the selection prescribed by the Schwartz inequality, that is, z ¼ ɑ again with > 0 and compute the corresponding value of the objective function. By letting ! 1, we realize that we converge to kɑk. This suggests that, for sufficiently large , we can approach the desired upper bound arbitrarily close (but there is no finite z for which we can attain it exactly!). This concludes the proof.

REFERENCES

[1] S. Periaswamy and H. Farid, "Elastic Registration in the Presence of Intensity Variation," IEEE Trans. Medical Imaging, vol. 22, no. 7, pp. 865-874, 2003.

[2] I. Karybali, E.Z. Psarakis, K. Berberidis, and G.D. Evangelidis, "Efficient Image Registration with Subpixel Accuracy," Proc. 14th European Signal Processing Conf., 2006.

[3] G.D. Hager and P.N. Belhumeur, "Efficient Region Tracking with Parametric Models of Geometry and Illumination," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 20, no. 10, pp. 1025-1039, Oct. 1998.

[4] J. Shi and C. Tomasi, "Good Features to Track," Proc. IEEE Int'l Conf Computer Vision and Pattern Recognition, 1994.

[5] M. Gleicher, "Projective Registration with Difference Decomposition," Proc.IEEE Int'l Conf. Computer Vision and Pattern Recognition, 1997.

[6] C. Fuh and P. Maragos, "Motion Displacement Estimation Using an Affine Model for Image Matching," Optical Eng., vol. 30, no. 7, pp. 881-887, 1991.

[7] Y. Altunbasak, R.M. Mersereau, and A.J. Patti, "A Fast Parametric Motion Estimation Algorithm with Illumination and Lens Distortion Correction," IEEE Trans. Image Processing, vol. 12, no. 4, pp. 395-408, 2003.

[8] B.K.P. Horn and E.J. Weldon, "Direct Methods for Recovering Motion,"Int'l J. Computer Vision, vol. 2, no. 1, pp. 51-76, 1988.

[9] P. Anandan, "A Computational Framework and an Algorithm for the Measurement of Visual Motion," Int'l J. Computer Vision, vol. 2, no. 3, pp. 283-310, 1989.

[10] B.D. Lucas and T. Kanade, "An Iterative Image Registration Technique with an Application to Stereo Vision," Proc. Seventh Int'l Joint Conf. Artificial Intelligence, 1981.

[11] E.Z. Psarakis and G.D. Evangelidis, "An Enhanced Correlation-Based Method for Stereo Correspondence with Sub-Pixel Accuracy," Proc. 10th IEEE Int'l Conf. Computer Vision, 2005.

[12] R. Szeliski, Handbook of Mathematical Models of Computer Vision, N. Paragios, Y. Chen, and O. Faugeras, eds., chapter 17. Springer, 2005.

[13] S. Baker and I. Matthews, "Lucas-Kanade 20 Years On: A Unifying Framework: Part 1. The Quantity Approximated, the Warp Update Rule, and the Gradient Descent Approximation," Int'l J. Computer Vision, vol. 56, no. 3, pp. 221-255, 2004.

[14] M.J. Black and Y. Yacoob, "Tracking and Recognizing Rigid and Non-Rigid Facial Motions Using Local Parametric Models of Image Motion," Proc. Fifth IEEE Int'l Conf. Computer Vision, 1995.

[15] H. Shum and R. Szeliski, "Construction of Panoramic Image Mosaics with Global and Local Alignment," Int'l J. Computer Vision, vol. 36, no. 2, pp. 101130, 2000.

[16] S. Nagahdaripour and C.H. Yu, "A Generalized Brightness Change Model for Computing Optical Flow," Proc. Fourth IEEE Int'l Conf. Computer Vision, 1993.

[17] B.K.P. Horn and B.G. Schunk, "Determining Optical Flow," Artificial Intelligence, vol. 17, pp. 185-203, 1981.

[18] M.J. Black and P. Anandan, "A Framework for the Robust Estimation of Optical Flow," Proc. Fourth IEEE Int'l Conf. Computer Vision, 1993.

