A Contrast Invariant Approach to Motion Estimation. Validation and Application to Motion Estimation Improvement

Vicent Caselles[1], Luis Garrido[1], and Laura Igual[1]

Universitat Pompeu Fabra, P. Circumval.lació, 8, Barcelona (Spain)
{vicent.caselles, luis.garrido, laura.igual}@upf.edu

Summary. We consider a contrast invariant approach to motion estimation which uses the direction of the gradient fields. The approach is region-based and assumes an affine motion model for each region. We propose to check if the estimated motion parameters fit properly the apparent motion of the region by a motion significance analysis. Moreover, we propose a motion field improvement which consider those regions that are not properly estimated according to the significance analysis and reassign them a motion model of a properly estimated neighboring region.

1 Introduction

Most known motion estimation methods employ the *intensity constancy assumption*, however, global or local illumination changes may violate this assumption and prevent the correct motion to be estimated. In [1] a constraint based on *spatial gradient's constancy* is proposed [2]. The *direction of the spatial gradient* is invariant with respect to global light changes and is insensitive to changes in illumination direction [3]. The work presented in [4] is based on the last property. The contrast invariance is incorporated in our approach [5] by the assumption that the shapes of the image move along the sequence.

In this work we present a hypothesis testing analysis approach that allows to measure how well the motion has been correctly estimated. This measure is used to validate the estimated motion parameters. Moreover, the validation output is used to reassign to the not properly estimated regions a motion model of a neighboring region which has been properly estimated.

The paper is organized as follows: Sect. 2 summarizes our motion estimation approach and presents the motion validation approach, Sect. 3 introduces the motion model reassignment approach, Sect. 4 presents some results and Sect. 5 ends up with the conclusions and future research work.

2 Motion Estimation and Significance Analysis

Region-based Contrast Invariant Motion Estimation: Let $I\colon \Omega \to R$ be a given image, where Ω is the image domain. The shapes of the image are identified with the family of its level lines which is a contrast invariant geometric description of the image [6]. The main assumption is they move along the image sequence (with possible deformation).

Motion is estimated between two frames of the sequence, denoted by I_0 and I_1, and $\phi(\mathbf{x})$ denotes the coordinates of the point at image I_1 whose coordinates are $\mathbf{x}$ at image I_0. Using the unit normals to describe the level lines, we propose to compute the optical flow ϕ by aligning the unit normal vector field $Z^1(\mathbf{x})$ to the level lines of I_1 with the transformed vector field of $Z(\mathbf{x})$ by the map ϕ, denoted by $\overline{Z}_\phi = (D\phi)^\dagger Z / \|(D\phi)^\dagger Z\|$ if $(D\phi)^\dagger Z \neq 0$ and 0 otherwise. $(D\phi)^\dagger$ denotes the cofactor matrix associated to $D\phi$, see [5].

Moreover we follow a region-based strategy, assuming that the motion fields can be expressed locally by a six parameter affine model. Let $\mathcal{R}$ be a partition into disjoint connected regions of the image I_0 bounded by level lines. The partition may be computed for instance with a segmentation algorithm like the Mumford-Shah functional subordinated to the topographic map [7].

Motion is estimated by minimizing the energy functional:

$$E_{\mathcal{R}}(\phi) := \frac{1}{\mathcal{N}_P} \sum_{j=1}^{\mathcal{N}_P} E_{R_j}(\phi) := \frac{1}{\mathcal{N}_P} \sum_{j=1}^{\mathcal{N}_P} \frac{1}{N_{R_j}} \sum_{\mathbf{x} \in R_j} \Psi \left(\frac{1}{4} \left\| Z^1(\phi(\mathbf{x})) - \overline{Z}_\phi(\mathbf{x}) \right\|^2 \right)$$

(1)

where $\Psi(.)$ may represent a robust function, the factor $1/4$ is used to normalize the cost term to the range $[0, 1]$, $\mathcal{N}_P = \mathrm{card}(\mathcal{R})$, and $N_{R_j} = \mathrm{card}(\{x \in R_j\})$. Motion is estimated using a gradient descent technique applied over E_R for each region R. For more details on this issue we refer to [5].

Motion Significance analysis: The minimization of E_R to estimate the motion parameters ϕ_R for any particular region R will always find a certain minimum, be it local or global. We cannot ensure that such minimum corresponds to the correct motion. Our purpose is to give a measure of the degree in which the motion has been correctly estimated. Let $R \in \mathcal{R}$ be a region of the image I_0. We consider the following hypotheses (which will be interpreted below): H_0: "the motion field of R is correct" and H_1: "the motion field of R is not correct".

Hypothesis Testing: Given a statistical model of the population, the observed sample is analyzed in order to see if it can be explained by it. If the observation diverges too much from the statistical model, the observation is rejected as belonging to the population.

For each region R, let ϕ_R the estimated motion parameters, and $\{\mathbf{x}_i\}_R = \{\mathbf{x}_i \in R \,/\, \|\nabla I_1(\phi(\mathbf{x}_i))\| > \gamma\}$, and $L = \mathrm{card}(\{\mathbf{x}_i\}_R)$. The threshold γ is used to ensure that the gradient orientations are not much affected by the presence of noise. We assume that the points in $\{\mathbf{x}_i\}_R$ are "independent" [8]. For each $\mathbf{x}_i \in \{\mathbf{x}_i\}_R$ we consider the unitary vectors $\mathcal{Z}^1(\phi(\mathbf{x_i}))$ as a random variable

and thus $\mathcal{Y}^i(\phi) = \Psi(1/4 \left\| \mathcal{Z}^1(\phi(\mathbf{x}_i)) - \overline{\mathcal{Z}}_\phi(\mathbf{x_i}) \right\|^2)$ may be interpreted as a random variable measuring the alignment of the two normal vectors. As in [8], we may consider that the vectors $\mathcal{Z}^1(\phi(\mathbf{x_i}))$ and $\overline{\mathcal{Z}}_\phi(\mathbf{x_i})$ are not aligned if they form an angle larger than a given threshold. We define the random variable $\mathcal{E}_R = \frac{1}{L} \sum_{i=1}^{L} \rho\left(\mathcal{Y}^i(\phi)\right)$, where $\rho : [0, +\infty[\to [0, +\infty[$ is an increasing function. Since $\mathcal{Y}^i(\phi)$ is directly related to the angle forming the two unitary vectors, its non-alignment can be subsumed into $\mathcal{E}_R$ by taking $\rho(x) := \rho_\beta(x) = 1$ if $x > \beta$, and 0 otherwise. In that case, $\mathcal{E}_R$ is a measure of the number of non-alignments for a given region R. We denote $E_R(\phi_R)$ the observed value of $\mathcal{E}_R$ corresponding to the data. The motion field ϕ_R assigned to R is correct if the error $E_R(\phi_R)$ is "sufficiently" small. Our purpose in this work is to define the region of rejection or acceptance using probability theory. If we assume that H_0 is true, the rejection region is of the form $[\mathcal{E}_R \geq \delta]$, $\delta > 0$ [9], but instead of computing the value of δ for a given level of significance as is usually done in hypothesis testing, we compute the probability $P[\mathcal{E}_R \geq E_R(\phi_R)|H_0]$ which corresponds to the probability of *miss-detection* or error of type I (to reject H_0 erroneously) for the observed value $\delta = E_R(\phi_R)$.

The probability that at least k_0 non-alignments occur is given by the binomial tail: $P[\mathcal{E}_R \geq k_0/L] = \mathcal{B}(p_0, k_0, L)$, where the probability of non-alignment p_0 is computed from the empirical data. Thus, the validation can be based in the expected number of miss-detections which is defined as follow.

The number of miss-detections (NMD) of a region is defined as $\mathrm{NMD}(R, \phi) = \mathcal{N}_P \cdot P[\mathcal{E}_R \geq E_R(\phi_R)|H_0]$, where $\mathcal{N}_P$ is the number of tested regions.

For a given region R, we reject H_0 if $\mathrm{NMD}(R, \phi) < \epsilon_0$. In that case we say that the motion of R is not properly estimated. Assuming that the motion model has been correctly estimated, the differences $\left\| \mathcal{Z}^1(\phi(\mathbf{x_i})) - \overline{\mathcal{Z}}_\phi(\mathbf{x_i}) \right\|$ (and therefore also $\mathcal{Y}^i(\phi)$) should be interpreted as noise. The probability p_0 is computed by $p_0 \approx \sum_{j=1}^{\mathcal{N}_P} \sum_{i=1}^{L_j} \rho_\beta(Y^i(\phi_{R_j}))/\sum_{j=1}^{\mathcal{N}_P} L_j$, where $L_j = \mathrm{card}(\{\mathbf{x}_i\}_{R_j})$ and $Y^i = \Psi(1/4 \left\| Z^1(\phi(\mathbf{x}_i)) - \overline{Z}_\phi(\mathbf{x}_i) \right\|^2)$.

A Contrario Model: The a contratio models were introduced in [8] as a tool for Gestalt's detection. In this context, as it is usual in this type of approach, we check that our observations cannot be explained by random selection of the motion model, that is, the number of coincidences between both vector fields $\mathcal{Z}^1(\phi(x))$ and $\overline{\mathcal{Z}}_\phi(x)$ is too large to be explained by a fortuite coincidence. In this case, we reject H_1 and we consider the motion model to be validated. For the precise details of this approach, we refer to [10].

3 Application: Motion Reassignment

An interesting application of the previous validation analysis is the enhancement of the overall motion estimation field. We propose to reassign to the not properly estimated regions the motion models of neighboring regions which have been correctly estimated. Neighboring regions may belong to the same
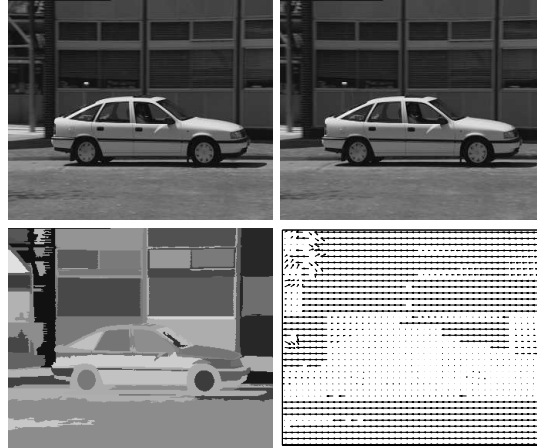
Fig. 1. Region-based motion example. From left to right and top to bottom. Original frame I_0, original frame I_1, partition of original frame I_0, recovered motion field.
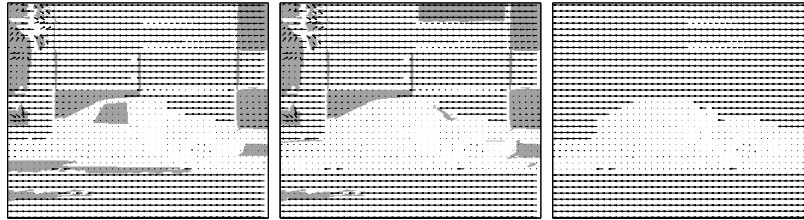

Fig. 2. Validation and reassignment example. From left to right. Validation using hypothesis testing and the a contrario model. Motion reassignment for the Vectra sequence.

moving object and thus they may have similar motions. The $\mathrm{NMD}(R, \phi)$ may be used to compare the different motion models that can be assigned to R. The higher the NMD is the better does the motion model explain the region movement. For each not properly estimated region the neighboring motion model leading to the highest NMD is assigned.

4 Results

In all experiments below $\Psi(r) = \sqrt{r^2}$, two vectors are aligned if they form an angle below 16° and $\epsilon_0 = 0.1$ for the statistical models.

Fig. 1 (top) shows frames #9 and #10 of the *vectra* sequence. The apparent motion of the car is zero while the background translates from right to left. The partition and the resulting motion vector field are shown in Fig. 1 (bottom). Some regions in the boundaries of the image and others near the car have

an incorrect estimated motion field. A validity process becomes necessary to detect these errors. Fig. 2 displays the outcome of each of the strategies presented in Sect. 2. Regions that are found as wrongly (resp. well) estimated are grey-shaded (resp. white). We have set the modulus gradient threshold γ to 7. Note that the validation strategies have been able to detect most of the wrongly estimated regions. It can be seen that the a contrario model is more demanding than the hypothesis testing one. The a contrario model can be considered a validation method, whereas the hypothesis testing model only performs an error control. The obtained motion field after the reassignment can be seen in Fig. 2 (right). We can identify different coherent moving regions of the scenes in these motion field.

5 Conclusions

Following [5], we interpret the image sequence as a set of moving level lines and we compute the optical flow as generated by a deformation between the level lines of two consecutive frames. We have introduced a motion significance measure based on hypothesis testing. These measures are useful both to detect the possible motion estimation errors and as a basic criterion for motion reassignment. Some issues have to be further developed in the future: automatic selection of the modulus gradient threshold γ, and the sensibility of the validation analysis to motion bias with respect the correct motion.

References

1. Bertero, M., Poggio, T.A., Torre, V.: Ill-posed problems in early vision. Procedings of the IEEE **76** (1988) 869–889
2. Papenberg, N., Bruhn, A., Brox, T., Didas, S., Weickert, J.: Highly accurate optic flow computation with theoretically justified warping. IJCV (2005)
3. Chen, H., Belhumeur, P., Jacobs, D.: In search of illumination invariants. In: ICCVPR. (2000) 254–261
4. Burgi, P.Y.: Motion estimation based on the direction of intensity gradient. IVC **22** (2004) 637–653
5. Caselles, V., Garrido, L., Igual, L.: A Contrast Invariant Approach to Motion Estimation. In: Int. Conf. on Scale Space, Springer Verlag (2005)
6. Caselles, V., Coll, B., Morel, J.M.: Topographic maps and local contrast changes in natural images. IJCV **33** (1999) 5–27
7. Ballester, C., Caselles, V., Igual, L.: Minimal morphological shape selection for segmentation and encoding. Preprint (2005)
8. Desolneux, A., Moisan, L., Morel, J.M.: Maximal Meaningful Events and Applications to Image Analysis. The Annals of Statistics **31** (2003) 1822–1851
9. Rohatgi, V.K., Saleh, M.E.: An Introduction to Probability and Statistics. Wiley-Interscience (2001)
10. Igual, L.: Image Segmentation and Compression using The Tree of Shapes of an Image. Motion Estimation. PhD thesis, Universitat Pompeu Fabra (2005)