Rotational Invariant Operators based on Steerable Filter Banks

X. Shi
Apple Computer, Inc.
xshi@apple.com

A.L. Ribeiro Castro
Dept. of Mathematics
UC Santa Cruz
alcastro@ucsc.edu

R. Manduchi
Dept. of Comp. Eng.
UC Santa Cruz
manduchi@soe.ucsc.edu

R. Montgomery
Dept. of Mathematics
UC Santa Cruz
rmont@count.ucsc.edu

Abstract

We introduce a technique for designing rotation invariant operators based on steerable filter banks. Steerable filters are widely used in Computer Vision as local descriptors for texture analysis. Rotation invariance has been shown to improve texture-based classification in certain contexts. Our approach to invariance is based on solving the PDE associated with the formulation of invariance in a Lie group framework.

I. Introduction

Image texture analysis is normally performed by first extracting feature vectors (*descriptors*), designed so as to capture the most relevant visual information. The descriptors, and the subsequent analysis algorithm, should be invariant under possible "nuisance" parameters. For example, when there is no reason to assume that the scene is viewed at a particular orientation, it is desirable that the response of the analysis be invariant under image rotation.

There are four main approaches in the literature for achieving rotation invariance. The first one is to add randomly rotated versions of the training samples when learning a statistical model or a classifier [4], [17], [19], [18]. Another method is to find the dominant orientation and normalize the descriptor vector with respect to it [10], [1], [9], [15], [12]. A third one is to only use descriptors that are naturally invariant under rotation, such as differential invariants [2], [3], [11], [16], integral invariants (moments) [8], [22], [21], [1], or circular symmetric filter kernels [24]. Finally, one may implement an operator that transforms non–invariant descriptors into invariant ones [6], [20]. This last approach is particularly attractive, since it does not restrict the choice of texture descriptors to a particular class (any texture descriptor can, in principle, be made rotation invariant), and does not require particular training procedures.

A general theory for the design of invariant operators was proposed by Van Gool at al. in [23]. The idea is to model the effect of nuisance parameters (in our case, rotations) as orbits of Lie group actions. This theoretical framework directly yields the number of independent invariants, as well as a systematic approach for finding them. The goal of this paper is to determine a maximal set of rotational invariants for a widely used class of texture descriptors, the so–called steerable filters. We show that, using the kernels proposed in [5], the Lie group theory of [23] can be used to find $N-1$ independent rotational invariant descriptors starting from the original N steerable filters. This requires solving a first order linear PDE. A closed form solution for this case is provided in this paper.

II. Problem Statement and Solution

A. Steerable Filters as Local Descriptors

A mainstream approach to texture analysis is based on the representation of texture patches by low–dimensional local descriptors. A linear filter bank is normally used for this purpose. The vector (descriptor) formed by the outputs of N filters at a certain pixel is a rank–N linear mapping of the graylevel profile within a neighborhood of that pixel. The marginal or joint statistics of the descriptor are then used to characterize the variability of texture appearance.

It is customary to choose analysis filters that are scaled and rotated version of one or more prototype kernels (with odd and even symmetry). If the prototype filter kernel is well localized in both spatial and frequency domains, this approach provides an effective sampling of the local spectrum along the semantically meaningful axes of scale and orientation. This representation is also attractive because it transforms in a predictable way under the action of similarity transformations of the domain (isotropic scaling, in–plane rotation). For a given kernel scale σ, let $\ell(x, y, \theta)$ be the descriptor component at pixel (x, y) corresponding to the kernel at orientation θ. One may expect that, after rotation of the image around (x, y) by angle $\Delta\theta$, the new output should be approximately equal to $\ell(x, y, \theta - \Delta\theta)$.

Rotation invariance has often been advocated for texture analysis [6], [24]. The quest for invariance stems from the fact that the orientation of the camera in the scene or of a surface patch with respect to the camera cannot be constrained (generic viewpoint assumption). A rotationally invariant operator transforms a texture feature into a quantity that is independent of any rotation of the camera around its optical axis.

We will concentrate on rotation invariant operators for a specific class of filter banks, the so–called *steerable filters* [5]. In its most general form, steerability is defined as the property of a function to transform under the action of a Lie group as a linear combination of a set of fixed *basis functions* [7]. We will consider here only steerable filters whose basis functions are the rotated versions of a prototype kernel $h(x, y)$ (*equivariant*

function space [7]). We will also assume that the prototype kernel for the steerable filter bank is axially symmetric, so that only rotations between 0 and π are of interest. If $h_i(x, y)$, with $0 \leq i < N$, is the version of the kernel rotated by $i\pi/N$, then the rotated version of $h(x, y)$ by angle θ can be written as a linear combination of the $h_i(x, y)$, with coefficients that only depend on θ. This also implies that, if $\ell_i \triangleq \ell(x, y, \theta_i)$ is the filtered version of an image $\ell(x, y)$ rotated by $i\pi/N$, then, for a generic angle θ:

$$\ell(x, y, \theta) = \sum_{i=0}^{N-1} \ell_i(x, y) k_i(\theta) \tag{1}$$

for suitable *interpolation functions* $k_i(\theta)$ that are independent of ℓ. We will concentrate on steerable prototype kernels that are higher order directional derivatives (in x) of an isotropic Gaussian function $G(x, y)$. Such functions can be written as $(-1)^M H_M(x) G(x, y)$, where $H_M(x)$ is an Hermite polynomial of order M, and M is the order of the derivative. It was shown in [5] that $N = M + 1$ basis functions are sufficient to synthesize $H_M(x) G(x, y)$. It was also shown in [5] that the interpolation functions in this case are trigonometric polynomials:

$$k_i(\theta) = \frac{2}{M+1} \sum_{n=0}^{(M-1)/2} \cos(2n+1)(\theta - \theta_i) \quad \text{(odd } M\text{)}$$

$$k_i(\theta) = \frac{1}{M+1} \left(1 + 2 \sum_{n=1}^{M/2} \cos 2n(\theta - \theta_i) \right) \quad \text{(even } M\text{)}$$

The next subsections will show how to impose rotation invariance to a steerable filter bank.

B. Invariant Operators and Lie Groups

Van Gool et al. [23] showed that designing invariants is equivalent to finding the solutions of a system of partial differential equations (PDE), thereby providing a systematic algorithm for design. Assume that for a given pixel (x, y), the descriptor vector is $(\ell_0, \ldots, \ell_{N-1})$. Since the group of planar rotations, *SO(2)*, is one-dimensional, we expect that $N - 1$ independent rotational invariants exist. More precisely, we define a "rotationally invariant operator" to be a mapping $f = (f_0, \ldots, f_{N-2})$ from the N–dimensional vector space representing the descriptors to an $(N - 1)$–dimensional manifold, whose components f_m solve the following PDE:

$$\sum_{i=0}^{N-2} \frac{\partial f_m}{\partial \ell_i} \frac{\partial \ell_i}{\partial \theta} \bigg|_{\theta=0} = 0 \tag{2}$$

where $\frac{\partial \ell_i}{\partial \theta}\big|_{\theta=0}$ represents the rate of change of the descriptor component ℓ_i as the image is rotated around pixel (x, y). Of course, the trivial invariant (f_m constant) solves (2), but is of no practical interest. We are looking for *maximal rank* descriptors [14], that is, a family of descriptors such that the rank of their Jacobian matrix is $N - 1$ almost everywhere. It is useful to note that any function $\Phi(f_0, \ldots, f_{N-2})$ is also rotationally invariant, i.e. solves the PDE above.

Thus, in order to find a rotationally invariant operator, we should be able to (a) express the partial derivatives of the measurements with respect to rotation, and (b) solve the PDE in (2). In Sec. II-C we show that step (a) is easily accomplished for steerable filter banks that are based on Gaussian derivatives. The solution to the PDE for this case is described in Sec. II-D.
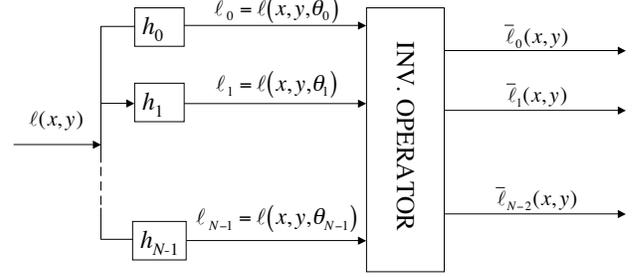

Fig. 1. The block scheme of a steerable filter bank with added rotation invariant operator.

C. Rotation Invariant Steerable Filters

A block scheme of the steerable filter bank with the addition of an invariant operator, transforming the vector $(\ell_0, \ldots, \ell_{N-1})$ into the rotation invariant vector $(\bar{\ell}_0, \ldots, \bar{\ell}_{N-2})$ where $\bar{\ell}_m = f_m(\ell_0, \ldots, \ell_{N-1})$ and $(f_0, \ldots f_{N-2})$ are the components of the rotationally invariant operator, is shown in Fig. 1. It is easy to prove that, thanks to the equivariant nature of the chosen filter bank, the PDE in (2) can be expressed in simple form as:

$$\nabla f_m^T H \underline{\ell} = 0 \tag{3}$$

where $\nabla f_m = (\partial f_m/\partial \ell_0, \ldots, \partial f_m/\partial \ell_{N-1})^T$, $\underline{\ell} = (\ell_0, \ldots, \ell_{N-1})^T$, and H is a Toeplitz antisymmetric matrix. In Lie theoretic terms, H is the "infinitesimal generator" of the action of the rotation group on the ℓ's. Tab. 1 reports the matrix H for filter order M from 1 to 3 (remember that $N = M + 1$).

D. Solving the PDE

We use the method of characteristics and linear algebra to solve the PDE (3). The matrix H of the PDE encodes the action of the circle group as per equation (2) and (1). Write $\underline{\ell}(\theta)$ for the vector $\underline{\ell}$ rotated by θ, and $d\ell/d\theta$ for its derivative with respect to θ. Then $\underline{\dot{\ell}} = \frac{1}{\omega} H \underline{\ell}$, where ω is a constant representing angular speed ($\omega = d\theta/dt$) and the superscript dot signifies differentiation with respect to t. Rewriting this ordinary differential equation (ODE) in terms of the independent variable t with $\theta = \omega t$ we obtain: $\underline{\dot{\ell}} = H \underline{\ell}$. This ODE is called the "characteristic equation" of our PDE. The solution curves for the ODE are the rotates of a given $\underline{\ell}$. The PDE (2) asserts the constancy of the function f along

$M = 1$
$H = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$
$f_0 = \ell_0^2 + \ell_1^2$

$M = 2$
$H = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}$
$f_0 = \ell_0 + \ell_1 + \ell_2$ $f_1 = \ell_0^2 + \ell_1^2 + \ell_2^2$

$M = 3$
$H = \begin{pmatrix} 0 & -1 & \frac{1}{\sqrt{2}} & -1 \\ 1 & 0 & -1 & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & 1 & 0 & -1 \\ 1 & \frac{-1}{\sqrt{2}} & 1 & 0 \end{pmatrix}$
$f_0 = \ell_0^2 + \ell_1^2 + \ell_2^2 + \ell_3^2$ $f_1 = \ell_0\ell_1 - \ell_0\ell_3 + \ell_1\ell_2 + \ell_2\ell_3$ $f_2 = -(\ell_0^4 + \ell_1^4 + \ell_2^4 + \ell_3^4) + 6(\ell_0^2\ell_2^2 + \ell_1^2\ell_3^2)$ $-\sqrt{2}((\ell_1 + \ell_3)\ell_2(\ell_2^2 - 3\ell_0^2) + (\ell_2 - \ell_0)\ell_3(\ell_3^2 - 3\ell_1^2)$ $+(\ell_1 - \ell_3)\ell_0(\ell_0^2 - 3\ell_2^2) + (\ell_2 + \ell_0)\ell_1(\ell_1^2 - 3\ell_3^2))$

TABLE I

THE MATRIX H AND A SET OF INDEPENDENT INVARIANTS $\{f_i\}$ FOR DIFFERENT STEERABLE FILTER ORDERS M.

these solution curves: $f(\underline{\ell}(\theta)) = f(\underline{\ell})$. To solve the PDE we first solve the ODE.

In the following, we will concentrate on the case $M = 3$ (see Tab. 1). To solve the ODE in this case, use a change of basis matrix B to put H in a block diagonal form: $J = B^{-1}HB$ with:

$$J = \begin{pmatrix} 0 & -\omega_1 & 0 & 0 \\ \omega_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\omega_2 \\ 0 & 0 & \omega_2 & 0 \end{pmatrix}$$

We compute:

$$B = \begin{pmatrix} 0 & -\sqrt{2} & 0 & -\sqrt{2} \\ 1 & -1 & 1 & 1 \\ \sqrt{2} & 0 & -\sqrt{2} & 0 \\ 1 & 1 & 1 & -1 \end{pmatrix}$$

Write $\underline{y} = B^{-1}\underline{\ell}$. In the y-variables the ODE becomes $\underline{\dot{y}} = J\underline{y}$. Upon setting: $z_1 = y_1 + iy_2$, $z_2 = y_3 + iy_4$ where $i = \sqrt{-1}$, the y-ODE becomes: $\dot{z}_1 = i\omega_1 z_1$, $\dot{z}_2 = i\omega_2 z_2$ which is a system of uncoupled ODEs. Its general solution is: $z_1(t) = e^{i\omega_1 t}z_1$, $z_2(t) = e^{i\omega_2 t}z_2$ where (z_1, z_2) now represent initial conditions. The frequencies $\pm i\omega_1, \pm i\omega_2$ are the eigenvalues of H. We compute that: $\omega_1 = \omega$, $\omega_2 = 3\omega$, $\omega = \sqrt{2}/2$. The fact that (ω_1, ω_2) is a multiple of an integer vector (here $(1, 3)$) is no accident. This must happen because H came from a circle action. The value of the multiplier ω has the interpretation of $\omega = d\theta/dt$ mentioned earlier.

The solutions to our PDE (3) (see (2)) are precisely the functions invariant under this flow. Clearly the functions $f_0 = |z_1|^2$ and $f_1 = |z_2|^2$ are invariant. Because of the $1 : 3$ resonance in the frequencies, the (complex) function $z_1^3 z_2^*$ is also invariant (where z^* denotes the complex conjugate of the

complex number z). Take its real part as the third invariant f_2. These functions will exhaust the invariants, i.e. the set of all solutions, algebraically. In other words, any smooth solution to our PDE is of the form $f = \Phi(f_0, f_1, f_2)$. In real terms: $f_0 = y_1^2 + y_2^2, f_1 = y_3^2 + y_4^2$, $f_2 = y_1^3 y_3^3 - 3y_1 y_2^2 y_3 + 3y_1^2 y_2 y_4 - y_2^3 y_4$.

Reverting to the old variables ℓ_i we compute:

$$f_0 = \frac{1}{8}(\ell_0^2 + \ell_1^2 + \ell_2^2 + \ell_3^2 + \sqrt{2}\ell_1(\ell_0 + \ell_2) + \sqrt{2}\ell_3(\ell_2 - \ell_0));$$

$$f_1 = \frac{1}{8}(\ell_0^2 + \ell_1^2 + \ell_2^2 + \ell_3^2 - \sqrt{2}\ell_1(\ell_0 + \ell_2) + \sqrt{2}\ell_3(\ell_0 - \ell_2));$$

$$f_2 = \frac{1}{64}(-\ell_0^4 - \ell_1^4 - \ell_2^4 - \ell_3^4 - \sqrt{2}(\ell_0^3(\ell_1 - \ell_3) + \ell_2^3(\ell_1 + \ell_3))$$
$$- \sqrt{2}\ell_2(\ell_1^3 + \ell_3^3) + 3\ell_3^2(2\ell_1^2 + \sqrt{2}\ell_1\ell_2) + 3\sqrt{2}\ell_1^2\ell_2\ell_3$$
$$- \sqrt{2}\ell_0(\ell_1 - \ell_3)(\ell_1^2 - 3\ell_2^2 + 4\ell_1\ell_3 + \ell_3^2)$$
$$+ 3\ell_0^2\ell_2(2\ell_2 + \sqrt{2}(\ell_1 + \ell_3)))$$
$$= \frac{1}{64}(-(\ell_0^4 + \ell_1^4 + \ell_2^4 + \ell_3^4) + 6(\ell_0^2\ell_2^2 + \ell_1^2\ell_3^2)$$
$$- \sqrt{2}((\ell_1 + \ell_3)\ell_2(\ell_2^2 - 3\ell_0^2) + (\ell_2 - \ell_0)\ell_3(\ell_3^2 - 3\ell_1^2) +$$
$$(\ell_1 - \ell_3)\ell_0(\ell_0^2 - 3\ell_2^2) + (\ell_2 + \ell_0)\ell_1(\ell_1^2 - 3\ell_3^2)));$$

Simplified forms for the first two functions can be obtained as $f_0 + f_1$ and $f_0 - f_1$ (reported in Tab. 1). Note that, for $M=1$, the found invariant is the squared magnitude of the smoothed gradient. This could have been obtained directly by using first order gauge coordinates.

Fig. 2 shows contour plots of the rotational invariant operators f_0, f_1 and f_2 applied to the kernels $h_i(x, y)$ for $M=3$. It should be noted that the rotational invariant operators are non–linear, and therefore the impulse responses *does not* fully characterize the system.
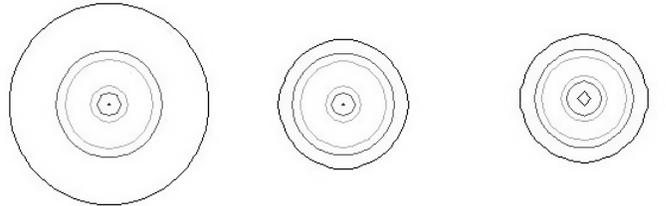

Fig. 2. Contour plots of the rotational invariant operators f_0, f_1 and f_2 applied to the kernels $h_i(x, y)$ for $M=3$. Larger values are represented with lighter grey.

Generalization to higher N. The method above will work for any N. There will be a change of basis (invertible matrix) B so that $J = B^{-1}HB$ is in block-diagonal form with 2×2 blocks of the form: $\begin{pmatrix} 0 & -\omega k_j \\ \omega k_j & 0 \end{pmatrix}$ with k_j integers and ω fixed. If N is odd there will be one additional 0 block. In the transformed variables $\underline{y} = B^{-1}\underline{\ell}$ the ODE decouples as before.

III. Discussion

The method just described is an application of the theory of representations of groups. To summarize this approach, let us reformulate the problem at hand. Equation (1) asserts that the ℓ_i span a finite-dimensional vector space V on which the circle group – the group of rotations of the plane – acts linearly. The quotient space of V by this action is by definition the set whose 'points' are orbits of the circle action. For example, a single point in the quotient consists of all rotates of ℓ_1. The map which sends a point in V to the orbit through that point is called the quotient map. The map f we are looking for is a realization of this quotient map. In concrete terms, the components f_i of the quotient map will consist of a basis of invariant polynomials. (A polynomial $p = p(\ell_1, \ldots, \ell_N)$ on V is called 'invariant' if its values do not change when all the ℓ_i are simultaneously rotated according to the circle action.) These f_i are the same f_i appearing in equation (2). So our problem is to find a basis of invariant polynomials for the given representation.

Representation theory tells us how to decompose V into "irreducible subspaces". An irreducible subspace for V is a linear subspace $P_j \subset V$ which is mapped to itself under the circle action, but which contains no subspace with this property. For the circle group, all irreducible subspaces have dimension 1 or 2. A two-dimensional irreducible subspace P_j is a copy of the usual plane, and rotation by θ acts on this plane like rotation by $k_j\theta$ does on the usual plane. Here k_j is an integer depending on P_j. In matrix terms this means that we can find a change of basis $y_i = \sum_j C_i^j \ell_j$ from the ℓ_i where the new basis y_i has the following property. Set $z_j = y_{2j-1} + iy_{2j}$. Then image rotation by θ has the effect $z_j \mapsto exp(ik_j\theta)z_j$. (The plane P_j is the span of y_{2j-1}, y_{2j}. If $N = 2k+1$ is odd then the last variable y_{2k+1} is unchanged by rotations and corresponds to a one-dimensional irreducible subspace.) It is now a simple matter for us to write down the desired invariants f_m, i.e. the solution to our problem, in terms of these diagonalizing variables. These invariants are the real quadratic functions $|z_j|^2$, together with the real and imaginary parts of those (complex) monomials $z_1^{\alpha_1} z_2^{\alpha_2} \ldots z_j^{\alpha_j}$ whose integer exponents α_i satisfy $k_1\alpha_1 + k_2\alpha_2 + \ldots k_j\alpha_j = 0$. If $\alpha_j < 0$ we must interpret $z_j^{\alpha_j}$ to mean $z_j^{*-\alpha_j}$, and not $1/z_j^{|\alpha_j|}$.

References

[1] Y.S. Abu-Mustafa and D. Psaltis, "Image normalization by complex moments", *IEEE Trans. PAMI*, 7(1), Jan. 1985.

[2] P.J. Besl and R.C. Jain, "Three-dimensional object recognition", *ACM Computing Surveys*, 17(1), 1985.

[3] P.E. Danielsson, "Rotation-invariant linear operators with directional response", *5th Intl Conf. Patt. Rec.*, Miami, Dec. 1980.

[4] H. Drucker, R. Schapire, and P. Simard, "Boosting performances in neural networks", *Int. Journal of Pattern Recogn. and Artif. Intell.*, 7:705–719, 1993.

[5] W.T. Freeman and E.H. Adelson, "The design and use of steerable filters", *IEEE Trans. PAMI*, 13(9):891–906, 1991.

[6] H. Greenspan, S. Belongie, R. Goodman, P. Perona, S. Rakshit, and C. H. Anderson, "Overcomplete steerable pyramid filters and rotation invariance.", *IEEE CVPR*, 1994.

[7] Y. Hel-Or and P. Teo, "Canonical Decomposition of Steerable Functions," *IEEE CVPR*, 1996.

[8] M-K. Hu, "Visual pattern recognition by moment invariants", *IRE Trans. on Information Theory*, IT-8:179-187, 1962.

[9] M. Kass and A. Witkin, "Analyzing oriented patterns", *Comp. Vision Graphics Image Proc.*, 37:362–385, 1987.

[10] H. Knutsson and G.H. Granlund, "Texture analysis using two-dimensional quadrature filters", *IEEE Workshop on Comp. Arch. Patt. Anal. Image Database Mgmt.*, 388–397, 1983.

[11] J.J. Koenderink and A.J. van Doorn, "Representation of local geometry in the visual system, *Biological Cybernetics*, 55:367375, 1987.

[12] D.G. Lowe, "Distinctive image features from scale-invariant keypoints", *International Journal of Computer Vision*, 60(2):91–110, November 2004.

[13] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors", *IEEE CVPR*, 2003.

[14] P.J. Olver, *Applications of Lie groups to differential equations*, Springer–Verlag, 1993.

[15] P. Perona and J. Malik, "Detecting and localizing edges composed of steps, peaks and roofs", *IEEE ICCV*, 1990.

[16] C. Schmid and R. Mohr, "Local grayvalue invariants for image retrieval", *IEEE Trans. PAMI*, 5(19):530–4, May 1997.

[17] B. Schölkopf, C. Burges, and V. Vapnik, "Incorporating invariance in support vector learning machines", *Artificial Neural Networks – ICANN'96*, 1996.

[18] X. Shi and R. Manduchi, "On rotational invariance for texture recognition", *4th International Workshop on Texture Analysis and Synthesis*, Beijing, China, 2005.

[19] P. Simard, B. Victorri, Y. Le Cun, and J. Denker, "Tangent prop – a formalism for specifying selected invariances in an adaptive network", *Adv. NIPS 4*, 1992.

[20] E. P. Simoncelli, "A rotation-invariant pattern signature", *IEEE Int'l Conf on Image Processing*, Lausanne, Switzerland, 1996.

[21] M.R. Teague, "Image analysis via the general theory of moments", *J. Opt. Soc. Am.*, 70:920–930, Aug. 1980.

[22] C.Teh and R.T. Chin, "On image analysis by the method of moments", *IEEE Trans. PAMI*, 10(4):496-513, 1988.

[23] L. Van Gool, T. Moons, E. Pauwels and A. Ooserlinck, "Vision and Lie's approach to invariance", *Image and Vision Computing*, 13(4):259–77, 1995.

[24] M. Varma and A. Zisserman, "A statistical approach to texture classification from single images", *IJCV*, 2005.