

CONTOUR-BASED OBJECT TRACKING WITH GRADIENT-BASED CONTOUR ATTRACTION FIELD

Petr Dokladal, Raffi Enficiaud and Eva Dejnozkova

School of Mines of Paris, Center of Mathematical Morphology, 35, Rue Saint Honoré,
77 300 Fontainebleau, FRANCE, e-mail: {dokladal,enficiau,dejnozke}@cmm.ensmp.fr

ABSTRACT

Object tracking is a problem which is very often addressed by various approaches as the background subtraction, visual constancy or geometric flow, to cite only a few.

The contribution of this paper is twofold: First, it proposes an enhanced feature-weighted gradient, enhancing the contours of the object characterized by this feature, and attenuating the gradient amplitude wherever the feature is not found.

The second, and main contribution, is a new integrator for active-contours based tracking. It uses the gradient alongside the object contours and generates a bidirectional attraction field.

The combination of the contour tracking and the weighted gradient is illustrated on the driver's-face tracking problem, where the illumination conditions (brightness, colour) change as the vehicle passes through different environments.

1. CAR DRIVER'S FACE TRACKING CONTEXT

The drivers' drowsiness is a non negligible problem tied to the road security. It can be evaluated by measuring the speed of the eyelids which must be completely non-invasive. The driver's face localization is often the first step to extract the eyes. Truck drivers' face detection has already been studied under simplified conditions, with among other things, the presence of a black colored curtain in the background. However, the context of personal vehicles has no such limiting conditions, and we have observed virtually no constant feature in the sequences.

Face detection with a reasonable reliability has been done under more general purpose conditions with template matching, feature extraction, colour detection or recently with more complex, training requiring techniques such as PCA, ICA or SVM, [1] or neural networks. A recent survey of various techniques can be found in [2]. Indeed, these techniques are less useful for face localization. For practical reasons (reduction of the amount of data, scale-independence), they usually operate on progressively subsampled images, and report a boolean flag indicating the presence of a face.

A growing amount of applications use colour to achieve specific tasks, explained both by its high separability properties, and the efficient tools improving classic segmentation tasks it provides. The skin hue, which is an inter-ethnic constant, seems to be a fairly good cue to initialize higher-level processes. Several way to represent colour as digital information exist, each defining a colour space endowed with particular properties. Among them, perceptual ([3]) and cylindrical ones provide tools for measuring colour differences. Besides, the way colours are represented in these spaces are strongly influenced by the illuminating light taken as reference ([4]). The difficulty stands in the fact that no supposition can be made about the conditions under which the images were obtained. Thus, important changes in incident light's brightness and chrominance avoid the use of skin clue solely, which reliability rapidly decreases as soon as the detection conditions deteriorate. A comparative study of performances of various colour spaces concerning this issues can be found in [5], [6] or [7].

Indeed, in the context of personal vehicles, the conditions are extremely rude, as the illumination is directional, the images' acquiring system itself is often saturated due to a strong contrast induced by the back light. Moreover, the light conditions rapidly change due to the vehicle motion, and usually a strong motion field is observed in the background.

The organization of the paper is as follow. Section 2 presents examples of gradients used in various colour spaces. We also introduce a feature-weighted gradient enhancing contours of objects characterized by a given feature and weakening those where the feature is not found. Section 3 introduces a new bidirectional attraction force used for contour tracking. Finally, section 4 shows how can the contour-tracking be combined with the feature-weighted gradient to extract car driver's face in video sequences, taken in personal vehicle under real conditions.

2. COLOUR GRADIENT

Although the notion of colour is intuitive, the way it is introduced into digital computation is rather complex. As men-

tioned above, colour provides high separability properties, and even if colour is difficult to use in an absolute manner, i.e. for modeling purposes, several tools measuring the differences between two colours exist.

Consider in the following the family of applications from set M to set N , noted $\mathcal{A}(M, N)$, and $\mathcal{I} \in \mathcal{A}(\mathcal{Z}^2, \Omega)$, a discrete image with value into the colour space Ω . In order to compare two colours in the image's colour space, a distance function on Ω , $d_\Omega \in \mathcal{A}(\Omega^2, \mathcal{R}_+)$, is required. Although the set of distance functions is infinite in each color space, some suggest a preferred metric. This is the case for $L^*a^*b^*$ and L^*u^*v , in which colour differences are given by euclidean distance based on norm \mathcal{L}_2 . For cylindrical colour spaces, such as HLS, angular opening d^α is used for the hue coordinate, and the distance based on norm \mathcal{L}_1 for the other coordinates.

Once a distance d defined on Ω , we are able to define gradient operator, noted ∇_Ω^d , along with this specific distance function for images mapping to Ω . The gradient operator is a spatial operator providing the edges of the regions or objects featured in the picture on which it is applied. Concerning the spatial properties of the gradient, different metrics can be used, and among them, with $N(p)$ being the neighborhood of point p (for instance the 4-connected neighborhood):

$$\begin{aligned}\nabla_\Omega \mathcal{I}(p) &= \max_{p_i \in N(p)} d_\Omega(\mathcal{I}(p), \mathcal{I}(p_i)) \\ \nabla_\Omega \mathcal{I}(p) &= \max_{p_i \in N(p)} d_\Omega(\mathcal{I}(p), \mathcal{I}(p_i)) \\ &\quad - \min_{p_i \in N(p)} d_\Omega(\mathcal{I}(p), \mathcal{I}(p_i))\end{aligned}$$

Compound operators, combining two or more previously defined gradients, were also proposed ([8]), such as the following one:

$$\nabla_{HLS} \mathcal{I}(p) = x_S \nabla^{d^\alpha} \mathcal{I}_H(p) + (1 - x_S) \nabla^{d^{\mathcal{L}_1}} \mathcal{I}_L(p)$$

computed on $\Omega \equiv HLS$ color space, \mathcal{I} being an image with values into $[0, 2\pi] \times [0, 1]^2$, \mathcal{I}_L and \mathcal{I}_H the projection of \mathcal{I} respectively into luminance and hue channels, and x_S the saturation's value of \mathcal{I} at point p . This gradient gives more weight either to hue or to luminance, inside regions of respectively high or poor saturation.

The computation of an image's gradient is usually the first step for more complex processes such as segmentation. Inside the obtained gradient map, the higher is the value of a given point, the stronger is the difference found in neighborhood of this point. However, one of the drawbacks of an operator defined this way is that it cannot include any specific a priori information. If we want to stress a particular feature of the image or just weaken an unneeded feature, we need to introduce a modified gradient operator.

The feature is related to a particular space Ω' which is not necessarily the same as the one used for the gradient's

computation. Consider an image function $\mathcal{I}' \in \mathcal{A}(\mathcal{Z}^2, \Omega')$, defined on the same support of \mathcal{I} . The feature can be expressed as a subset $\mathcal{F} \in \Omega'$. For instance, if the feature is the likeness of a point for being a skin point, then Ω' can be *HLS* color space, which is a colour space providing accurate ability to differentiate skin points from non-skin points ([9]). Another example may be the motion field on which the velocity of point represents the feature, in this case $\Omega' \equiv \mathcal{R}^2$, and one may search for points with a certain velocity.

We define a distance function to the feature \mathcal{F} as follow:

$$d_{\Omega'|\mathcal{F}}(x) = d_{\Omega'}(x, \mathcal{F}) = \min_{y \in \mathcal{F}} d_{\Omega'}(x, y), x \in \Omega'$$

We eventually define the modified gradient operator $\nabla_F : \mathcal{A}(\mathcal{Z}^2, \Omega) \rightarrow \mathcal{A}(\mathcal{Z}^2, \mathcal{R}_+)$ as follow:

$$\nabla_F \mathcal{I}(p) = \frac{\nabla_\Omega \mathcal{I}(p)}{1 + d_{\Omega'|\mathcal{F}}(\mathcal{I}'(p))}, p \in \mathcal{Z}^2 \quad (1)$$

This modified gradient keeps the edges of skin areas intact, and smooths the edges of regions that are of no interest with respect to \mathcal{F} . More generally, this gradient allows to keep particular features intact while suppressing the others, so that the convergence of algorithms focusing on these features will be eased.

3. OBJECT TRACKING

Object tracking is a frequently studied topic. Various methods using different approaches have been proposed, based either on i) *visual constancy* which uses some inherent feature of the object, e.g. the colour, ii) *geometric flow* which extracts the motion field or iii) the *background subtraction* which often requires some a priori knowledge about the background. In general, some limiting assumptions must be made about the scene, such as strong contrast alongside the contours, homogeneity, static background, etc. Besides these model-free methods, one may employ some object model, trying to fit the data to a given model.

Active contours (or snakes) adapt to the form of the data. In the next section we present a new integrator to to implement a model-free, contour based, adaptive contour segmentation in the terms of level sets.

3.1. A gradient-based attraction field

Consider an image \mathcal{I} and some gradient of \mathcal{I} , $g = \nabla \mathcal{I}$. Let

$$g_K = g * K \quad (2)$$

where K is some triangular window: $\mathcal{Z}^2 \rightarrow \mathcal{R}^+$, such that

$$K(x, y) = \begin{cases} 1 - \alpha(x^2+y^2)^{1/2} & \text{if } (x^2+y^2)^{1/2} < 1/\alpha \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

Note that in the signal processing domain, convoluting with such a window is a frequency filter. However, filtering is not the objective here.

∇g_K represents a gradient-dependent integrator with interesting properties. Generally, the evolution of a curve \mathcal{C}

$$\frac{\partial \mathcal{C}}{\partial t} = F \vec{n} \quad (4)$$

where \vec{n} is the normal vector to \mathcal{C} , and F represents the motion speed. For the contour-based tracking we propose

$$F = \nabla g_K \quad (5)$$

It can be shown (by approximating g in Eq. 2 by a Dirac impulse δ , and computing F in Eq. 5 in a discrete form) that ∇g_K is a bidirectional integrator pointing toward the crest of the gradient g from both sides.

The advantage of using bidirectional integrator is twofold: i) it allows the contour to converge toward the gradient maximum from both sides, and ii) it eliminates the necessity to use a constant one-directional attraction force there, where the data is zero. This fact eliminates the problem of local breaches in the gradient, often introducing leakage in object reconstruction. Attempts to alleviate this problem were made in [10], introducing a viscous watershed, capable to slow down the propagation in such narrow openings. Although the leakage could probably be alleviated by using curvature, the leakage problem does not occur when using ∇g_K , since on zero gradient the contour does not move.

This fact is also used in defining the capture range of the contours, which is in fact limited. Suppose that the maximum inter-frame displacement of the object is bounded by D . This information should be taken into account by letting $\text{supp}\{(x, y) | K(x, y) > 0\}$ be a circle of radius D , generating a non-zero attraction field in a narrow zone around the contour. Hence, a convenient value of α in Eq. 3 is $a = 1/D$.

Indeed, as the attraction force stops on the zero crossing of the gradient, its principle is similar to the Haralick [11] edge detector, which detects edges on zero crossing of the second derivative of \mathcal{I} in the gradient direction. Kimmel in [12] reformulates the Haralick edge detector in terms of the level set framework and shows how it can be combined with additive constraints to segment images. As stated before, our objective is the contour-based object tracking. Whereas various motion predictors can be used to predict the displacement direction according to the past, arbitrary deformations of the object give birth to a displacement field with locally varying direction. Any contour-based tracking must therefore be able to handle both partially forward and backward displacement of the contour. A good overview of other existing attraction vector fields can be found in [13].

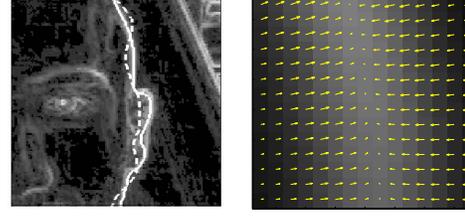


Fig. 1. (a) The initial (dashed) and final (solid line) position of the contour, and (b) zoom on the attraction force field F .



Fig. 2. (a) Original image, (b) usual euclidean gradient on Lab, (c) featured weighted gradient, the feature being the skin chroma.

4. APPLICATION

By using Eq. 4, the current contour C^n of the object is obtained by using the attraction field g_K^n generated by the current frame \mathcal{I}^n , and the contour C^{n-1} in the previous frame:

$$C^n = \lim_{T \rightarrow \infty} \int_0^T \nabla g_K^n(C) \vec{n} dt + C^{n-1}$$

with $C(t=0) = C^{n-1}$

where $g_K^n = g * K$,

$$g(p) = \frac{\nabla_{\text{Lab}} I(p)}{1 + d_{\text{HLS}|\mathcal{F}}(I(p))}$$

We use the feature based on the skin chroma. We take $\Omega' \equiv \text{HLS}$, and $\mathcal{F} = \{x \in \text{HLS} | x_H \in [-20^\circ, 50^\circ]\}$. This feature is only related to hue, thus the distance d_{HLS} used is the angular distance d^α to the skin chroma \mathcal{F} . The size of the triangular window K is ten pixels, i.e. $\alpha = 0.1$, calculated from a natural gesture speed as seen by our camera.

Initialization: The description of the initialization of the tracking is outside the scope of this paper. It can be successfully done by combining several features, see e.g. [14], using the face colour and shape or [15] combining the colour and motion (in a car application, no perturbing motion is present in the background before the car runs).

5. CONCLUSIONS

The contribution of this paper is twofold. First, it introduces a feature-weighted gradient, useful for enhancing the



Fig. 3. Contour tracking applied to driver's face extraction, using the weighted gradient (skin chroma being the feature of interest). Randomly chosen images from a video sequence.

contours of objects characterized by this feature, and weakening those where the feature is not found in the image.

Secondly, this paper proposes a gradient-based attraction field for model-less tracking of objects. This attraction field is generated by a bi-directional, data-dependent integrator.

The application section shows the combination of the contour-tracking and the feature-weighted gradient to track a car-driver's face, the feature being the skin hue. Although the skin hue is, rather an unstable feature due to varying illumination conditions, the gradient weighted by distance to the standard skin hue, remains more robust.

Perspectives: One of the interesting properties of the above-proposed attraction field is that it is zero on zero data (alleviating thus leakage on narrow breaches introduced by noise). If used in real applications, this attraction force can be combined with region regularizers, some a priori model, or curvature, for example.

6. REFERENCES

- [1] O. Déniz, M. Castrillón, and M. Hernández, "Face recognition using independent component analysis and support vector machines," *Pattern Recognition Letters*, , no. 24, pp. 2153–2157, 2003.
- [2] M.-H. Yang, D. J. Kriegman, and N. Ahuja, "Detecting faces in images: A survey," *Pattern Analysis and Machine Intelligence*, to appear available online: "<http://vision.ai.uiuc.edu/mhyang/face-detection-survey.html>".
- [3] Wysecki & Stiles, *Color Science : Concepts and Quantitative Data and Formulae*, John Wiley & Sons, New-York, 2nd edition, 1982.
- [4] Available online, "<http://www.brucelindbloom.com/>".
- [5] J.-C. Terrillon and S. Akamatsu, "Comparative performance of different chrominant spaces for color segmentation and detection of human faces in complex scene images," in *Proc. of the 12-th Conf. on Vision Interface*, 5 1999, vol. 2, pp. 180–187.
- [6] J. B. Martinkauppi, M. N. Soriano, and M. H. Laaksonen, "Behavior of skin color under varying illumination seen by different cameras at different color spaces," .
- [7] C. H. Lee, J. S. Kim, and K. H. Park, "Automatic face location in a complex background using motion and color information," *Pattern Recognition*, vol. 11, no. 29, pp. 1877–1889, 1996.
- [8] J. Angulo and J. Serra, "Color segmentation by ordered merging," in *IEEE International Conference on Image Processing (ICIP'03)*, Barcelona, Spain, September 14 - 17 2003, IEEE, number D-/03/MM, pp. 1–4.
- [9] Jean-Christophe Terrillon, Hideo Fukamachi, Shigeru Akamatsu, and Mahdad N. Shirazi, "Comparative performance of different skin chrominance models and chrominance spaces for the automatic detection of human faces in color images," in *Fourth IEEE International Conference on Automatic Face and Gesture Recognition*, Grenoble, 2000.
- [10] F. Meyer and C. Vachier, "Image segmentation based on viscous flooding simulation," in *International Symposium on Mathematical Morphology*. Sydney, Australia, 2002.
- [11] R. Haralick, "Digital step edges from zero crossing of second directional derivatives," *IEEE Transactions on Pattern Recognition and Machine Intelligence*, no. 6, pp. 58–68, 1984.
- [12] R. Kimmel and A. Bruckstein, "On edge detection integration and geometric active contours," in *International Symposium on Mathematical Morphology*. vol. 2, Sydney, Australia, 2002.
- [13] C. Xu and J. L. Prince, "Snakes, shapes, and gradient vector flow," *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 359–369, march 1998.
- [14] K. Sobottka and I. Pittas, "Extraction of facial regions and features using color and shape information," in *International Conference on Pattern Recognition (ICPR'96)*, August 1996, vol. III, pp. C421–C425.
- [15] A. Nayak and S. Chaudhuri, "Self-induced color correction for skin tracking under varying illumination," in *IEEE International Conference on Image Processing*, September 2003.