

Compression d’images par triangulation géodésique anisotrope

Sébastien BOUGLEUX¹, Gabriel PEYRÉ², Laurent COHEN²

¹ Université de Caen / GREYC ² CEREMADE / Université Paris Dauphine

sebastien.bougleux@unicaen.fr, {peyre, cohen}@ceremade.dauphine.fr

Résumé – Ce papier propose une nouvelle méthode de compression d’images utilisant des triangulations de Delaunay géodésiques. Ces triangulations sont générées par un maillage géodésique progressif qui exploite l’anisotropie des images à l’aide d’une méthode d’échantillonnage par point le plus éloigné. Cet échantillonnage est calculé avec une distance géodésique anisotrope, ce qui impose aux triangles de Delaunay géodésiques à suivre la géométrie de l’image. Ces calculs géodésiques sont effectués à l’aide d’un Fast Marching Riemannien, qui met à jour de façon récursive la carte des distances aux points déjà échantillonnés. Une interpolation par splines linéaires sur cette triangulation permet d’approcher fidèlement les contours saillants ainsi que les singularités directionnelles dans les images. La compression est obtenue en codant à la fois les coefficients de l’approximation spline, ainsi que la déviation entre la triangulation géodésique et la triangulation de Delaunay Euclidienne. Des résultats numériques montrent que la prise en compte de l’anisotropie améliore l’approximation par rapport à une triangulation isotrope. Le codeur géodésique ainsi obtenu permet une amélioration par rapport aux codeurs en ondelettes comme JPEG-2000 pour des images géométriques.

Abstract – We propose a new image compression method based on geodesic Delaunay triangulations. Triangulations are generated by a progressive geodesic meshing algorithm which exploits the anisotropy of images through a farthest point sampling strategy. This seeding is performed according to anisotropic geodesic distances which force the anisotropic Delaunay triangles to follow the geometry of the image. Geodesic computations are performed using a Riemannian Fast Marching, which recursively updates the geodesic distance to the seed points. A linear spline approximation on this triangulation allows to approximate faithfully sharp edges and directional features in images. The compression is achieved by coding both the coefficients of the spline approximation and the deviation of the geodesic triangulation from an Euclidean Delaunay triangulation. Numerical results show that taking into account the anisotropy improves the approximation by isotropic triangulations of complex images. The resulting geodesic encoder competes well with wavelet-based encoders such as JPEG-2000 on geometric images.

La compression d’images nécessite l’extraction des singularités géométriques d’une image afin de minimiser la courbe de débit-distortion. La transformée en ondelettes est un outil populaire et a donné naissance au standard JPEG-2000 [10]. Le support isotrope des ondelettes les rend cependant sous-optimales pour représenter les contours ou les texture orientées.

Une autre classe de méthodes utilise des triangulations adaptées à la géométrie de l’image afin de compresser efficacement les contours. Le calcul d’une triangulation optimale est cependant NP-difficile [1], et il faut donc calculer des triangulations approchées à l’aide d’algorithmes gloutons. Une première classe de méthodes procède par décimation d’une triangulation de Delaunay Euclidienne dense, ce qui donne de bons résultats en approximation de surfaces [6] ainsi qu’en compression d’images [3, 4].

L’utilisation d’une triangulation de Delaunay Euclidienne est problématique dans les régions de forte anisotropie, où une triangulation optimale dévie fortement d’une triangulation Euclidienne. Les méthodes de raffinement de maillages permettent une adaptation anisotrope locale de la métrique, par exemple pour mailler un domaine plan [9], pour résoudre des EDP [7] ou pour approximer des fonctions régulières ou *cartoon* [11].

Contributions. Nous proposons dans cet article d’utiliser une

métrique anisotrope globale, qui permet une prise en compte à grande distance de l’anisotropie des images. Cette méthode reprend des techniques de calculs géodésiques anisotropes introduites récemment en vision par ordinateurs [2]. Nous introduisons un algorithme de raffinement de Delaunay anisotrope, qui étend les algorithmes de raffinement par modification locale de la métrique [9] à une véritable distance géodésique anisotrope. Ceci permet un meilleur respect des contours des images naturelles même lorsque l’on utilise peu de triangles, dans quel cas une anisotropie locale est insuffisante.

1 Approximation par triangulation et anisotropie

Approximation par triangulation. Il est possible d’approcher une image, vue comme une fonction $f \in L^2([0, 1]^2)$, à l’aide de m fonctions splines linéaires sur une triangulation $(\mathcal{V}, \mathcal{F})$, où $\mathcal{V} = \{v_i\}_{i=0}^{m-1} \subset [0, 1]^2$ sont les sommets et $\mathcal{F} \subset \{0, \dots, m-1\}^3$ sont les indices des faces. On définit $f_m = \sum_i a_i \varphi_i$ où φ_i est linéaire sur chaque face de \mathcal{F} et $\varphi_i(v_j) = 1$ si $i = j$ et 0 sinon. Les coefficients a_i sont calculés pour que f_m soit la projection orthogonale de f sur l’espace V engendré par les φ_i ,

soit minimisant $\|f - g\|$ pour $g \in V$. Ceci nécessite la résolution du système linéaire $\Phi a = F$, avec $\Phi_{i,j} = \langle \varphi_i, \varphi_j \rangle$ et $F_i = \langle f, \varphi_i \rangle$.

Afin de minimiser l'erreur $\|f - f_m\|$ par rapport à la triangulation, il est nécessaire d'optimiser la position des points \mathcal{V} ainsi que les connexions \mathcal{F} entre ces points. On doit ainsi placer plus de points près des zones de fort gradient de l'image, et utiliser des triangles fortement anisotropes, orientés selon la direction des contours. On peut montrer (voir par exemple [10]) pour une image cartoon, qui est de classe C^2 à l'extérieur de contours C^2 , qu'une telle optimisation de la triangulation permet d'obtenir une erreur de l'ordre de $\|f - f_m\| = O(m^{-1})$, ce qui est bien meilleur que l'approximation m -termes en on-deletes ou par triangulation isotrope, qui génère une erreur de l'ordre de $O(m^{-1/2})$. Dans ce cas, les longueurs $\Delta_1 \geq \Delta_2$ des côtés des triangles qui sont près des contours doivent vérifier $\Delta_2 \approx (\Delta_1)^2$. Nous proposons un algorithme efficace de triangulation adaptée à ce problème.

Métrie anisotrope. L'anisotropie souhaitée d'une triangulation $(\mathcal{V}, \mathcal{F})$ peut se représenter à l'aide d'un champ de tenseurs, qui correspond à la donnée d'une matrice $H(x) \in \mathbb{R}^{2 \times 2}$ définie positive en chaque point $x \in [0, 1]^2$. Chaque tenseur est diagonalisé

$$H(x) = \lambda_1(x)e_1(x)e_1(x)^T + \lambda_2(x)e_2(x)e_2(x)^T,$$

où les valeurs propres sont $0 < \lambda_1 \leq \lambda_2$ et e_1, e_2 sont deux champs orthogonaux de vecteurs propres (non-orientés).

Chaque triangle de $(\mathcal{V}, \mathcal{F})$ se conformant à H , et centré en x_0 , doit être inscrit dans l'ellipsoïde

$$(x_0 - x)^T H(x_0)(x_0 - x) \leq t_m,$$

pour une valeur de t_m qui dépend du nombre m de triangles. Les côtés des triangles sont ainsi alignés avec le champ $e_1(x_0)$, et les côtés du triangle se conforment au ratio Δ_1/Δ_2 qui vaut approximativement $\sqrt{\lambda_2(x_0)/\lambda_1(x_0)}$.

On peut montrer que pour une fonction de classe C^2 , ce ratio $\lambda_2(x_0)/\lambda_1(x_0)$ peut être pris égal au ratio $|\mu_2(x_0)/\mu_1(x_0)|$ des valeurs propres de la hessienne de f au point x , ce qui améliore la constante dans l'erreur d'approximation $\|f - f_m\| = O(m^{-1})$, voir [11]. Cette estimation est cependant assez instable et ne permet pas de traiter des images bruitées, des textures orientées, ou des images cartoon dont les contours ne sont pas lissés.

Tenseur de structure. Nous utilisons donc un champ H calculé à partir du gradient de l'image, en définissant le tenseur de structure

$$T(x) = G_\sigma * (\nabla f(x)\nabla f(x)^T),$$

où G_σ est un filtre Gaussien appliqué à chaque composante du tenseur $\nabla f(x)^T \nabla f(x)$ de rang 1, voir [8]. Le paramètre σ permet d'obtenir une estimation robuste en présence de bruit, et nous avons choisi $\sigma = 5/n$ pour une image de taille $n \times n$.

À partir de la diagonalisation du tenseur de structure,

$$T(x) = \mu_1(x)e_1(x)e_1(x)^T + \mu_2(x)e_2(x)e_2(x)^T,$$

où (e_1, e_2) sont deux champs orthogonaux de vecteurs propres de H , on définit le tenseur de la triangulation comme

$$H(x) = (\varepsilon + \mu_1(x))^\alpha e_1(x)e_1(x)^T + (\varepsilon + \mu_2(x))^\alpha e_2(x)e_2(x)^T,$$

où ε contrôle l'adaptativité isotrope de la triangulation (la variation de densité, voir [12]) et α l'adaptativité anisotrope. Une métrique avec un fort ε est presque constante, et une métrique avec un faible α est presque isotrope. Nous avons fixé $\varepsilon = 10^{-5}$ et l'influence du paramètre α est adaptée en fonction du type d'image à traiter.

2 Distance géodésique

La métrique $H(x_0)$ définit localement la taille des triangles autour de x_0 , qui doivent être de taille constante t_m pour la métrique modifiée

$$\|x_0 - x\|_{H(x_0)}^2 = (x_0 - x)^T H(x_0)(x_0 - x).$$

Afin de placer de façon uniforme des points pour cette métrique, il faut définir une distance globale $d_H(x, y)$ pour n'importe quel couple de points $x, y \in [0, 1]^2$.

La longueur d'une courbe $\gamma : [0, 1] \rightarrow \Omega$ suivant la métrique $H(x)$ est définie comme

$$L_H(\gamma) \stackrel{\text{def}}{=} \int_0^1 \|\gamma'(t)\|_{H(\gamma(t))} dt.$$

Une courbe γ passant par un point $\gamma(t) = x$ avec une vitesse $\gamma'(t)$ a une longueur locale plus faible si $\gamma'(t)$ est approximativement colinéaire à $e_1(x)$. La distance géodésique entre deux points $x, y \in [0, 1]^2$ est définie comme

$$d_H(x, y) = \min_{\gamma} L_H(\gamma) \quad (1)$$

où les chemins γ sont réguliers par morceaux avec $\gamma(0) = x$ et $\gamma(1) = y$. Une courbe γ qui satisfait $L_H(\gamma) = d_H(x, y)$ est une géodésique joignant x à y . Si $H(x) = h$ est constante sur l'image, alors $d_H(x, y) = \|x - y\|_h$.

Étant donné un ensemble de points $\mathcal{S} = (x_k)_k \subset [0, 1]^2$, on définit la carte d'action minimale $U_{\mathcal{S}}(x) = \min_k d_H(x, x_k)$. Un résultat théorique important est que pour un champ H régulier, $U_{\mathcal{S}}$ est l'unique solution de viscosité de l'équation d'Hamilton-Jacobi

$$\|\nabla U_{\mathcal{S}}(x)\|_{H(x)^{-1}} = 1 \quad \text{avec} \quad U_{\mathcal{S}}(x_k) = 0. \quad (2)$$

Cette carte peut donc être calculée en résolvant numériquement cette équation sur une grille régulière de $N = n \times n$ pixels. Dans le cas particulier d'une métrique isotrope $H(x) = \lambda(x)\text{Id}_2$, l'algorithme de Fast Marching permet de résoudre une discrétisation par différence finie en amont de cette équation en $O(N \log(N))$ opérations. Plusieurs extensions de cette méthode au cas d'une métrique $H(x)$ générique ont été proposées, voir par exemple [13].

3 Calcul de maillages géodésiques anisotropes

Algorithme de point le plus éloigné. Le calcul d'une triangulation $(\mathcal{V}, \mathcal{F})$ se conformant à un champ de tenseur $H(x)$ est un problème d'optimisation difficile. La contrainte de conformité sur \mathcal{V} impose que les points v_i soient distribués uniformément selon la distance géodésique d_H , c'est à dire que $d_H(v_i, v_j)$ est approximativement constant si $(i, j) \in \mathcal{F}$.

On peut réaliser approximativement cette contrainte en utilisant la méthode d'échantillonnage par point le plus éloigné [5]. Pour calculer cet échantillonnage, Peyré et Cohen [12] ont proposé un algorithme basé sur le Fast Marching, dans le contexte du remaillage géodésique isotrope.

Cette méthode s'étend sans difficulté au cas d'une métrique Riemannienne quelconque, comme le détaille l'algorithme 1. L'étape 3 nécessite uniquement une mise à jour locale de la carte de distances à l'intérieur de la région

$$\{x \setminus U_{\{q\}}(x) \leq U_{\mathcal{V}-\{q\}}(x)\}.$$

Ceci rend les calculs de plus en plus rapides au fur et à mesure que de nouveaux points sont ajoutés.

Algorithm 1: Algorithme de point le plus éloigné.

Initialisation : soit $\mathcal{V}^{(0)} = \{v_0\}$ un coin de l'image.;

for $k = 2, \dots, m$ **do**

1. Calcul de $v_k = \operatorname{argmax}_{q \in \Omega} U_{\mathcal{V}^{(k-1)}}(q)$;
2. Mise à jour $\mathcal{V}^{(k)} = \mathcal{V}^{(k-1)} \cup \{v_k\}$;
3. Mise à jour de $U_{\mathcal{V}^{(k)}}$ à partir de $U_{\mathcal{V}^{(k-1)}}$.

Renvoyer : $\mathcal{V} = \mathcal{V}^{(m)}$.

Triangulation de Delaunay géodésique. Une fois l'échantillonnage \mathcal{V} obtenu, on calcule une triangulation \mathcal{F} garantissant l'alignement des faces avec la direction d'anisotropie locale $e_1(x)$. Ceci est obtenu en considérant le graphe de Delaunay géodésique de \mathcal{V} , qui est le dual du diagramme de Voronoï anisotrope de \mathcal{V} obtenu avec la métrique H . Une arête (i, j) fait partie de ce graphe si les régions de Voronoï V_i et V_j se touchent, où

$$V_i = \{x \setminus \forall k \neq i, d_H(v_i, x) \leq d_H(v_k, x)\}.$$

Une difficulté est que le graphe obtenu n'est pas toujours une triangulation plane du domaine de l'image. On peut démontrer que c'est le cas pour une métrique H régulière lorsque suffisamment de points sont ajoutés.

Nous avons donc modifié l'algorithme 1 de point le plus éloigné afin de garantir que les bords de l'image fassent toujours partie de la triangulation, que les segments $[v_i, v_j]$ pour tout $(i, j) \in \mathcal{F}$ ne se croisent pas, et qu'il n'existe pas de point isolé. Ces modifications sont similaires à celles développées dans [9], et nécessitent, à certaines étapes de l'algorithme, de modifier le placement du point v_i pour maintenir les contraintes.

4 Approximation et compression d'images

Approximation d'images. L'approximation f_m obtenue est progressivement raffinée lorsque l'on accroît la valeur de m . La figure 1, à gauche, montre que pour l'approximation de l'image Cameraman, l'erreur $\|f - f_m\|$ est minimale pour $\alpha \approx 0.75$, et ceci pour une large gamme de m . Une étude expérimentale semble indiquer que $\alpha = 0.75$ est un choix acceptable sur un large ensemble d'images naturelles, et nous avons utilisé cette valeur dans la suite.

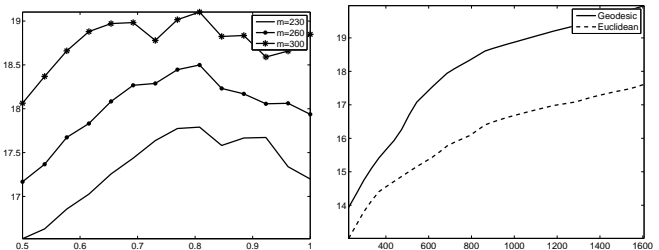


FIG. 1 – Gauche : pour $m = 230, 260, 300$, SNR en fonction de α . Droite : pour $\alpha = 0.75$, SNR en fonction de m pour des triangulations de Delaunay anisotropes (trait plein) et pour des triangulations de Delaunay Euclidiennes (pointillés).

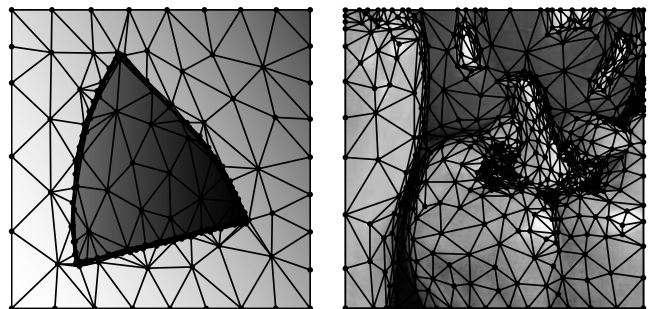


FIG. 2 – Exemples de triangulations calculées pour $m = 200$ (gauche) et $m = 600$ (droite).

La figure 2, à gauche, montre des exemples de triangulations pour deux valeurs de m et pour $\alpha = 0.75$. La figure 1, à droite, montre l'évolution, en fonction de m , du SNR pour une triangulation de Delaunay géodésique et Euclidienne de l'ensemble des points échantillonnés avec $\alpha = 0.75$. Ceci montre l'importance d'utiliser une connectivité adéquate en même temps qu'une triangulation anisotrope.

Compression d'images. Le codage de l'approximation f_m nécessite de quantifier à la fois la position des points v_i avec un pas T_v et les coefficients a_i de l'interpolation spline avec un pas T_a , pour obtenir des valeurs entières arrondies $\bar{v}_i = [v_i/T_v]$ et $\bar{a}_i = [a_i/T_a]$. Les coefficients \bar{a} et \bar{v} sont codés avec un codeur arithmétique à l'aide de $R_v + R_a$ bits. Le décodeur peut ainsi calculer des positions et valeurs décompressées

$\tilde{a}_i = (\bar{a}_i + \text{sign}(\bar{a}_i)/2)T_a$ et $\tilde{v}_i = (\bar{v}_i + \text{sign}(\bar{v}_i)/2)T_v$. On obtient alors une approximation $\tilde{f}_m = \sum_i \tilde{a}_i \tilde{\varphi}_i$, où $\tilde{\varphi}_i$ est la base spline associée aux positions \tilde{v}_i . Il est possible de remplacer le codeur arithmétique par un codeur contextuel plus efficace, ce qui diminuerait encore le nombre de bits $R_v + R_a$, voir par exemple [4].

La connectivité de la triangulation \mathcal{F} doit également être codée à l'aide de R_c bits. Afin de réduire R_c , nous calculons des modifications élémentaires pour passer de \mathcal{F} à la triangulation de Delaunay Euclidienne des points \tilde{v}_i , que le codeur comme le décodeur peuvent calculer. Étant donnée une arête (i, j) , incidente à deux triangles (i, j, k_1) et (i, j, k_2) qui ne sont pas dans $\tilde{\mathcal{F}}$, on remplace l'arête (i, j) par l'arête (k_1, k_2) . Le nombre de bits nécessaires pour coder \mathcal{F} est ainsi réduit à $R_c = K \log_2(A)$ où K est le nombre de remplacements et A le nombre d'arêtes dans \mathcal{F} .

D'autres compresseurs d'images exploitant des triangulations ont été proposés, comme par exemple [3, 4]. Ces approches partent d'une triangulation dense, qui est décimée afin de réduire l'erreur d'approximation. Notre approche est différente, car elle exécute un raffinement de maillage, ce qui est très rapide pour des petites valeurs de m . De plus, elle utilise une triangulation de Delaunay basée sur une distance géodésique anisotrope, où l'élongation des triangles s'adapte plus facilement aux structures de l'image comparativement à la triangulation de Delaunay Euclidienne utilisée par exemple dans [3].

Le choix de T_v et T_a peut être optimisé pour minimiser la courbe de débit-distortion $\|f - f_m\|$ en fonction du nombre de bits $R = R_v + R_a + R_c$. Nous avons utilisé par simplicité $R_v = 10^{-2}$ et $R_a = \|f\|_\infty/100$. La figure 3, gauche, compare des compressions obtenues à l'aide de notre méthode et avec JPEG-2000 (voir [10]). Ceci montre que pour de faibles débits, et pour des images suffisamment géométriques, notre méthode permet de gagner par rapport au standard actuel.

Références

- [1] P. Agarwal and S. Suri. Surface Approximation and Geometric Partitions. *SIAM J. Comput.*, 19 :1016–1035, 1998.
- [2] S. Bogleux, G. Peyre, and L.D. Cohen. Anisotropic geodesics for perceptual grouping and domain meshing. In *Proc. of ECCV'08*, pages II : 129–142, 2008.
- [3] L. Demaret, N. Dyn, and A. Iske. Image compression by linear splines over adaptive triangulations. *Signal Processing Journal*, 86(7) :1604–1616, 2006.
- [4] L. Demaret, A. Iske, and W. Khachabi. Contextual image compression from adaptive sparse data representations. In *Proc. of SPARS'09*, 2009.
- [5] Y. Eldar, M. Lindenbaum, M. Porat, and Y. Y. Zeevi. The farthest point strategy for progressive image sampling. *IEEE Trans. Image Processing*, 6(9) :1305–1315, 1997.
- [6] M. Garland and P. Heckbert. Surface simplification using quadric error metrics. *Proc. of SIGGRAPH 1997*, pages 209–215, 1997.

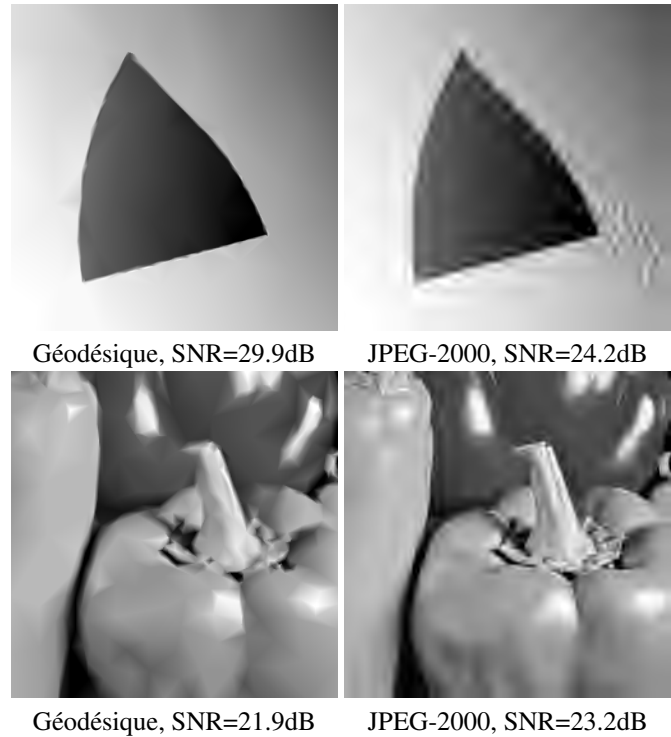


FIG. 3 – Gauche : Comparaison de notre schéma de compression avec JPEG-2000, pour le même nombre de bits, pour $m = 200$ (haut) et $m = 600$ (bas) triangles, pour $\alpha = 0.75$.

- [7] P.L. George, H. Borouchaki, P.J. Frey, P. Laug, and E. Saltel. Mesh generation and mesh adaptivity : theory, techniques. *Encyclopedia of computational mechanics*, E. Stein, R. de Borst and T.J.R. Hughes ed., John Wiley & Sons Ltd., 2004.
- [8] M. Kass and A. Witkin. Analyzing oriented patterns. *Comput. Vision Graph. Image Process.*, 37(3) :362–385, 1987.
- [9] F. Labelle and J. R. Shewchuk. Anisotropic voronoi diagrams and guaranteed-quality anisotropic mesh generation. In *19th Annual Symp. on Computational Geometry*, pages 191–200. ACM, 2003.
- [10] S. Mallat. *A Wavelet Tour of Signal Processing*, 3rd edition. Elsevier, 2009.
- [11] J-M. Mirebeau and A. Cohen. Greedy bisection generates optimally adapted triangulations. Technical report, Laboratoire Jacques-Louis Lions, 2008.
- [12] G. Peyré and L. D. Cohen. Geodesic remeshing using front propagation. *Int. Journal on Computer Vision*, 69(1) :145–156, 2006.
- [13] E. Prados, C. Lenglet, J. P. Pons, N. Wotawa, R. Deriche, O. D. Faugeras, and S. Soatto. Control theory and fast marching methods for brain connectivity mapping. In *Proc. CVPR'09*, volume 1, pages 1076–1083, 2006.