

Vision par ordinateur

Stéréoscopie par minimisation d'énergie

Frédéric Devernay
d'après le cours de Richard Szeliski

Mise en correspondance stéréo

Quels algorithmes possibles ?

- mettre en correspondance des "features" (points d'intérêts) et interpoler
- idem avec des contours
- mettre en corresp. des fenêtres de pixels

optimisation:

- mise à jour itérative
- programmation dynamique
- minimisation d'énergie (regularisation, stochastique)
- coupes de graphes

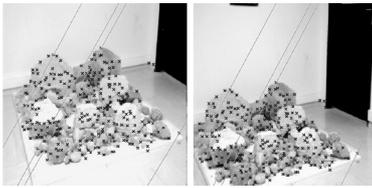
6/12/01

Stéréoscopie par minimisation d'énergie

2

Stéréo à base de points d'intérêt

Mise en correspondance de "coins" (pts d'intérêt)



Interpoler pour obtenir une carte de profondeur

6/12/01

Stéréoscopie par minimisation d'énergie

3

Interpolation de données

Étant donné un ensemble de pts 3D, comment *interpoler* pour obtenir une surface 3D ?

Interpolation de données éparées [Nielson93]

- trianguler dans une des deux images
- mettre dans une grille et remplir
- placer une *fonction de noyau* en chaque point
- minimiser une fonction d'énergie (membrane)

6/12/01

Stéréoscopie par minimisation d'énergie

4

Minimisation d'énergie

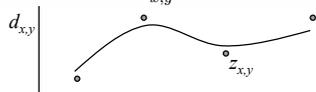
Exemple en 1D: splines d'approximation

$$E_{\text{total}}(\mathbf{d}) = E_{\text{data}}(\mathbf{d}) + \lambda E_{\text{smoothness}}(\mathbf{d})$$

$$E_{\text{data}}(\mathbf{d}) = \sum_{x,y} (d_{x,y} - z_{x,y})^2$$

$$E_{\text{membrane}}(\mathbf{d}) = \sum_{x,y} (d_{x,y} - d_{x-1,y})^2$$

$$E_{\text{thin plate}}(\mathbf{d}) = \sum_{x,y} (2d_{x,y} - d_{x-1,y} - d_{x+1,y})^2$$



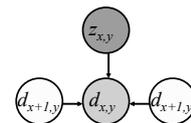
6/12/01

Stéréoscopie par minimisation d'énergie

5

Relaxation

Améliorer itérativement une solution en minimisant localement l'énergie : *relaxation*



Intuitif, implantation simple, mais vitesse de convergence lente

6/12/01

Stéréoscopie par minimisation d'énergie

6

Relaxation

Comment obtenir la formule de mise à jour ?
La différentielle de l'énergie doit être nulle

$$\frac{\partial E}{\partial d_{x,y}} = 2(d_{x,y} - z_{x,y}) + 2\lambda(2d_{x,y} - d_{x-1,y} - d_{x+1,y}) = 0$$

$$d_{x,y} \leftarrow \frac{1}{1+2\lambda}(z_{x,y} + d_{x-1,y} + d_{x+1,y})$$

6/12/01

Stéroréoscopie par minimisation d'énergie

7

Énergie non quadratique

Comment minimiser cette fonction de coût ?

$$E_{\text{total}}(\mathbf{d}) = E_{\text{data}}(\mathbf{d}) + \lambda E_{\text{smoothness}}(\mathbf{d})$$

$$E_{\text{data}}(\mathbf{d}) = \sum_{x,y} (d_{x,y} - z_{x,y})^2$$

$$E_{\text{smoothness}}(\mathbf{d}) = \sum_{x,y} |d_{x,y} - d_{x-1,y}|$$

$$\frac{\partial E}{\partial d_{x,y}} = 2(d_{x,y} - z_{x,y}) + \lambda[\text{sgn}(d_{x,y} - d_{x-1,y}) + \text{sgn}(d_{x,y} - d_{x+1,y})] = 0$$

6/12/01

Stéroréoscopie par minimisation d'énergie

8

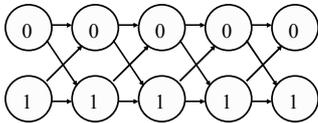
Éspace d'optimisation discret

Et si les valeurs de d sont discrètes ?

$$E_{\text{total}}(\mathbf{d}) = E_{\text{data}}(\mathbf{d}) + \lambda E_{\text{smoothness}}(\mathbf{d})$$

$$E_{\text{data}}(\mathbf{d}) = \sum_{x,y} (d_{x,y} - z_{x,y})^2$$

$$E_{\text{smoothness}}(\mathbf{d}) = \sum_{x,y} |d_{x,y} - d_{x-1,y}|$$



6/12/01

Stéroréoscopie par minimisation d'énergie

9

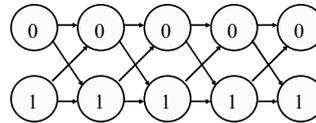
Programmation dynamique

Calculer le meilleur coût cumulé en chaque pixel

$$E_{\text{total}}(\mathbf{d}) = E_{\text{data}}(\mathbf{d}) + \lambda E_{\text{smoothness}}(\mathbf{d})$$

$$E_{\text{data}}(\mathbf{d}) = \sum_{x,y} (d_{x,y} - z_{x,y})^2$$

$$E_{\text{smoothness}}(\mathbf{d}) = \sum_{x,y} |d_{x,y} - d_{x-1,y}|$$



6/12/01

Stéroréoscopie par minimisation d'énergie

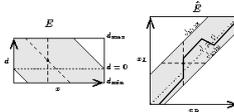
10

Programmation dynamique

Fonction de coût 1D

$$E(\mathbf{d}) = \sum_{x,y} \rho_P(d_{x+1,y} - d_{x,y}) + \sum_{x,y} E_0(x,y;d)$$

$$\tilde{E}(x,y,d) = E_0(x,y;d) + \min_{d'} (\tilde{E}(x-1,y,d') + \rho_P(d_{x,y} - d'_{x-1,y}))$$



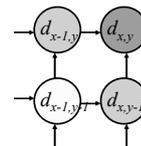
6/12/01

Stéroréoscopie par minimisation d'énergie

11

Programmation dynamique

Est-ce que ça fonctionne aussi en 2D ?



Non : $d_{x,y-1}$ et $d_{x-1,y}$ peuvent dépendre de valeurs différentes de $d_{x-1,y-1}$

6/12/01

Stéroréoscopie par minimisation d'énergie

12

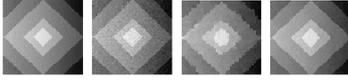
Coupes de graphe (graph cuts)

Technique de résolution d'un problème 2D

$$E_{\text{total}}(\mathbf{d}) = E_{\text{data}}(\mathbf{d}) + \lambda E_{\text{smoothness}}(\mathbf{d})$$

$$E_{\text{data}}(\mathbf{d}) = \sum_{x,y} f_{x,y}(d_{x,y})$$

$$E_{\text{smoothness}}(\mathbf{d}) = \sum_{x,y} \rho(d_{x,y} - d_{x-1,y}) + \sum_{x,y} \rho(d_{x,y} - d_{x,y-1})$$



(a) original image (b) observed image (c) local min w.r.t. standard moves (d) local min w.r.t. alpha-expansion moves

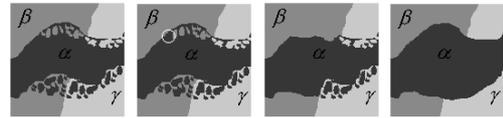
6/12/01

Stéréoscopie par minimisation d'énergie

13

Coupes de graphe (graph cuts)

Deux types de mouvements :



(a) initial labeling (b) standard move (c) alpha-beta-swap (d) alpha-expansion

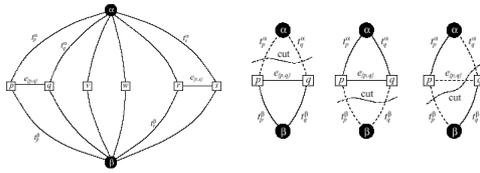
6/12/01

Stéréoscopie par minimisation d'énergie

14

Coupes de graphe (graph cuts)

alpha-beta swap: échanger les étiquettes alpha et beta



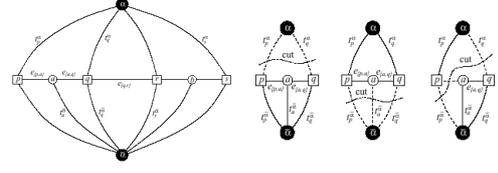
6/12/01

Stéréoscopie par minimisation d'énergie

15

Coupes de graphe (graph cuts)

alpha expansion: ajouter des pixels à la classe alpha



6/12/01

Stéréoscopie par minimisation d'énergie

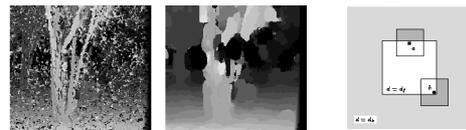
16

De retour à la stéréoscopie...

Taille de la fenêtre de corrélation

Petit voisinage : plus de détails

Grand voisinage : moins d'erreurs isolées



w = 3

w = 20

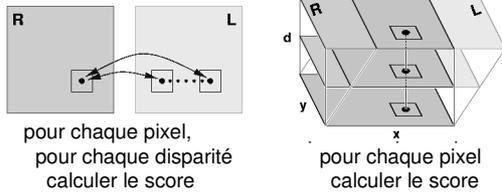
6/12/01

Stéréoscopie par minimisation d'énergie

18

Stéréoscopie par corrélation

Réordonnancement des boucles (pixel/disparité)



6/12/01

Stéréoscopie par minimisation d'énergie

19

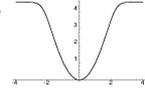
Stéréoscopie par corrélation

- Pour chaque disparité, calculer le coût en chaque pixel

$$E_0(x, y; d) = \rho(I_L(x' + d, y') - I_R(x', y'))$$

Pourquoi une fonction ρ robuste

- occlusions, réflexion spéculaires, autres...



On peut également utiliser un autre critère (cosinus)

6/12/01

Stéréoscopie par minimisation d'énergie

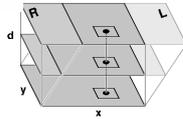
20

Stéréoscopie par corrélation

- Score = moyenne spatiale

$$E(x, y; d) = \sum_{(x', y') \in N(x, y)} E_0(x', y', d)$$

- On utilise une fenêtre carrée (facilement optimisable)



- On peut également utiliser une moyenne pondérée, de la diffusion [non-linéaire]...

6/12/01

Stéréoscopie par minimisation d'énergie

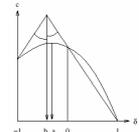
21

Stéréoscopie par corrélation

- Optimisation par fenêtre glissante : complexité indépendante de la taille de la fenêtre de corrélation
- En chaque pixel, meilleur score \rightarrow disparité
- Interpolation sous-pixélique parabolique ou "toît"

$$d = d_0 + \frac{1}{2} \frac{c(d_0 + 1) - c(d_0 - 1)}{(c(d_0) - c(d_0 + 1)) + (c(d_0) - c(d_0 - 1))}$$

$$d = d_0 + \begin{cases} \frac{1}{2} \frac{c(d_0 + 1) - c(d_0 - 1)}{c(d_0) - c(d_0 + 1)} & \text{si } c(d_0 + 1) > c(d_0 - 1), \\ \frac{1}{2} \frac{c(d_0 + 1) - c(d_0 - 1)}{c(d_0 + 1) - c(d_0)} & \text{si } c(d_0 + 1) < c(d_0 - 1). \end{cases}$$



- Validation par corrélation retour (échange des images)

6/12/01

Stéréoscopie par minimisation d'énergie

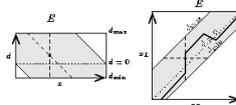
22

Programmation dynamique

Fonction de coût 1D [Intille & Bobick, IJCV 99]

$$E(\mathbf{d}) = \sum_{x,y} \rho_P(d_{x+1,y} - d_{x,y}) + \sum_{x,y} E_0(x, y; d)$$

$$\tilde{E}(x, y, d) = E_0(x, y; d) + \min_{d'} (\tilde{E}(x, y, d') + \rho_P(d_{x,y} - d'_{x,y}))$$



6/12/01

Stéréoscopie par minimisation d'énergie

23

Programmation dynamique

Image de l'espace de disparité et chemin optimal

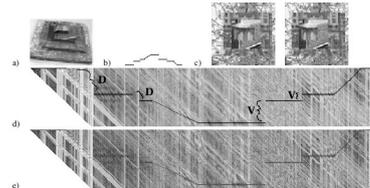


Fig. 4 This figure shows (a) a model of the stereo-slipping wedding cake that we will use as a test example, (b) a depth profile through the center of the slipping wedding cake, (c) a simulated, noise-free image pair of the cake, (d) the enhanced, cropped, correlation DSI for a stereo-slipping wedding cake (SNR = 10 dB), (e) (f) the regions labeled "V" mark diagonal pixels in the matching path caused by regions occluded in the left image. The regions labeled "V" mark vertical jumps in the path caused by regions occluded in the right image.

6/12/01

Stéréoscopie par minimisation d'énergie

24

Programmation dynamique

Image



Coupes de graphe



Programmation dynamique
(artefacts horiz.)



Fig. 12. Results of two stereo algorithms on Figure 1. (a) Original left image; (b) Cut of st. algorithm [46]; and (c) the algorithm described in this paper.

6/12/01

Stéréoscopie par minimisation
d'énergie

25

Coupes de graphe (graph cuts)

α - β swap

α expansion

la pénalité de lissage dépend de la présence de contours

calcul des meilleurs appariements pour des valeurs entières de la disparité

6/12/01

Stéréoscopie par minimisation
d'énergie

26

Inférence bayésienne

Formulation statistique du problème

Modèle a priori $p_P(\mathbf{d})$

Modèle de mesure $p_M(I_L, I_R | \mathbf{d})$

Modèle a posteriori (Th. de Bayes)

$$p_M(\mathbf{d} | I_L, I_R) \propto p_P(\mathbf{d}) p_M(I_L, I_R | \mathbf{d})$$

Maximum a posteriori (MAP) :

maximiser $p_M(\mathbf{d} | I_L, I_R)$

6/12/01

Stéréoscopie par minimisation
d'énergie

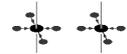
27

Champs de Markov

Distribution de probabilité du champ de disparité $d(x,y)$ ne dépend que des voisins

$$p_P(d_{x,y} | \mathbf{d}) = p_P(d_{x,y} | \{d_{x',y'}, (x',y') \in \mathcal{N}(x,y)\})$$

$$p_P(\mathbf{d}) = \frac{1}{Z_P} e^{-E_P(\mathbf{d})}$$



$$E_P(\mathbf{d}) = \sum_{x,y} \rho_P(d_{x+1,y} - d_{x,y}) + \rho_P(d_{x,y+1} - d_{x,y})$$

Permet de gérer le *lissage* ou la *cohérence* du champ

6/12/01

Stéréoscopie par minimisation
d'énergie

28

Modèle de mesure

Vraisemblance de la correspondance d'intensité

$$p_M(I_L, I_R | \mathbf{d}) = \frac{1}{Z_M} e^{-E_0(x,y;d)}$$

$$E_0(x,y;d) = \rho(I_L(x'+d,y') - I_R(x',y'))$$

Correspond à un bruit gaussien si ρ quadratique

6/12/01

Stéréoscopie par minimisation
d'énergie

29

Estimation MAP

Maximiser la vraisemblance à posteriori

$$E(\mathbf{d}) = -\log p(\mathbf{d} | I_L, I_R)$$

$$= \sum_{x,y} \rho_P(d_{x+1,y} - d_{x,y}) + \rho_P(d_{x,y+1} - d_{x,y})$$

$$+ \sum_{x,y} \rho_M(I_L(x+d_{x,y}, y) - I_R(x, y))$$

Équivalent à une *régularisation* (minimisation d'énergie avec contraintes de lissage)

6/12/01

Stéréoscopie par minimisation
d'énergie

30

Pourquoi l'estimation bayésienne ?

Principe déterminant la fonction de coût
Modèle explicite du bruit et de la connaissance a priori

Grande variété d'algorithmes d'optimisation :

- descente de gradient (minimisation locale)
- optimisation stochastique (échantillonnage de Gibbs)
- approximation du champ moyen
- théorie des graphes (déterministe) [Zabih]

6/12/01

Stéréoscopie par minimisation d'énergie

31

Approximation du champ moyen

Règle de diffusion non-linéaire bayésienne :

- mettre à jour la distribution de probabilité des disparités en supposant que les distributions des voisins sont indépendantes (vrai en 1D)

Équivalent à chercher la meilleure approximation *factorisée*

$$P(\mathbf{d}|I_L, I_R) \sim Q(\mathbf{d}) = \prod_i Q_i(d_i)$$

6/12/01

Stéréoscopie par minimisation d'énergie

32

Approximation du champ moyen

log estimée MAP

$$-\log P(\mathbf{d}|I_L, I_R) = \sum_{ij} E_{ij}(d_i, d_j) + \sum_i E_i(d_i) \\ = \sum_{ij} \mathbf{s}_i \mathbf{A}_{ij} \mathbf{s}_j + \sum_i \mathbf{b}_i \mathbf{s}_i$$

divergence de Kullback-Leibler

$$D_{KL} = H(Q) - \sum_{\mathbf{d}} Q(\mathbf{d}) \log P(\mathbf{d}) \\ = \sum_{ik} q_{ik} \log q_{ik} + \sum_{ij} \mathbf{s}_i \mathbf{A}_{ij} \mathbf{s}_j + \sum_i \mathbf{b}_i \mathbf{s}_i$$

6/12/01

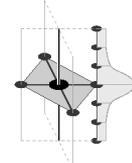
Stéréoscopie par minimisation d'énergie

33

Approximation du champ moyen

minimiser la divergence K-L avec

$$\sum_k q_{ik} = 1$$



règle de mise à jour :

$$q_{ik} \propto \exp[-(\sum_j \mathbf{a}_{ij}^k \mathbf{q}_j + \mathbf{b}_{ik})] \\ = \exp[-(\sum_{jl} E_{ij}(d_i=k, d_j=l) p(d_j=l) + E_i(d_i=k))]$$

6/12/01

Stéréoscopie par minimisation d'énergie

34

Cartes de disparité



Image d'entrée



Somme des différences absolues



Champ moyen



Coupes de graphe

6/12/01

Stéréoscopie par minimisation d'énergie

35

Stéréo avec diffusion non-linéaire

Avantages:

- fonctionne très bien dans dans régions sans occlusion

Inconvénients :

- limité à deux images (pas vraiment)
- se trompe dans les zones d'occlusion
- ne gère pas les *pixels mélangés*

6/12/01

Stéréoscopie par minimisation d'énergie

36

Bibliographie

- Y. Boykov, O. Veksler, and Ramin Zabih, Fast Approximate Energy Minimization via Graph Cuts, Unpublished manuscript, 2000.
- A.F. Bobick and S.S. Intille. Large occlusion stereo. *International Journal of Computer Vision*, 33(3), September 1999. pp. 181-200
- D. Scharstein and R. Szeliski. Stereo matching with nonlinear diffusion. *International Journal of Computer Vision*, 28(2):155-174, July 1998
- R. Szeliski. Stereo algorithms and representations for image-based rendering. In *British Machine Vision Conference (BMVC'99)*, volume 2, pages 314-328, Nottingham, England, September 1999.
- R. Szeliski and R. Zabih. An experimental comparison of stereo algorithms. In *International Workshop on Vision Algorithms*, pages 1-19, Kerkyra, Greece, September 1999.
- G. M. Nielson, Scattered Data Modeling, *IEEE Computer Graphics and Applications*, 13(1), January 1993, pp. 60-70.