

Sur Quelques Méthodes de Calcul de Flot Optique à partir du Tenseur de Structure : Synthèse et Contribution

About Some Optical Flow Methods from Structure Tensor : Review and Contribution

F.-B. Lauze¹

P. Kornprobst²

C. Lenglet²

R. Deriche²

M. Nielsen¹

¹ IT University of Copenhagen

² INRIA Sophia Antipolis, Projet Odyssée

Glentevej 67-69,
DK-2400 Copenhagen NV
Denmark
francois@it-c.dk

Résumé

Ce papier traite de l'estimation du flot optique à partir de l'analyse du champ de tenseurs de structure d'une séquence vue comme un volume spatio-temporel. Il a deux objectifs. Le premier est de présenter une synthèse de quelques approches récentes en établissant leur lien. Nous discutons en particulier les phases d'extraction du flot à partir d'un champ de tenseurs et de d'obtention de tenseurs lissés. Le second est de proposer une équation de diffusion non linéaire sur les tenseurs contrôlée par la mesure intuitive des coins. L'ensemble de ces approches fera l'objet d'expérimentations sur plusieurs séquences test.

Mots Clef

Flot optique, tenseur de structure, cohérence, diffusion non linéaire, équations aux dérivées partielles

Abstract

This paper is about optical flow estimation from structure tensor field analysis of a sequence seen as a spatio-temporal volume. It tackles two issues. The first is to review related recent techniques showing their relations. We essentially discuss the stages of flow extraction from structure tensor fields and the obtention of regularized tensor fields. The second is to propose a nonlinear approach controlled by the intuitive corner measure. The overall approaches will be compared in details on several test sequences.

Keywords

Optical flow, structure tensor, coherence, nonlinear diffusion, partial differential equations

1 Introduction

Une idée intuitive pour analyser une séquence d'images est de la considérer comme un volume spatio-temporel et d'analyser sa structure. La structure dans le volume regroupe l'information spatiale des images pour un temps donné mais aussi son évolution temporelle. Le problème de l'analyse de vitesses devient analogue à celui de l'analyse des orientations dans un volume 3D. Une méthode pour analyser cette structure est de calculer le Tenseur de Structure (TS) qui décrit une information locale moyennée des orientations. A l'origine, ce moyennage était implémenté au moyen de convolutions par des gaussiennes spatiales ou spatio-temporelles [17]. Si l'on voit la convolution gaussienne comme la solution d'une Equation aux Dérivées Partielles (EDP) de diffusion, on peut alors envisager d'appliquer des EDP linéaires ou non linéaires pour améliorer la qualité des résultats. A partir des nombreuses contributions dans le domaine des EDP appliquées au traitement d'images (voir par exemple [3, 32] pour des synthèses), des méthodes ont été proposées pour mieux diffuser la donnée ou le tenseur de structure pour en extraire le flot. Parmi ces approches, on peut citer [22, 25, 35, 7, 18, 22, 21, 20].

La démarche dans ce type d'approches comprend trois ingrédients. Le premier est une phase de prétraitement du volume par convolution ou diffusion non linéaire. Le second est le lissage du champ des tenseurs de structure de manière à récupérer au mieux l'information de vitesse. Le troisième est la phase d'extraction qui consiste à obtenir une estimation du flot à partir des tenseurs. Cette description parcellaire illustre les différents types de traitements mis en jeu, mais elle aussi caricaturale : certaines

approches combinent par exemple les lissages de l'image et des tenseurs. L'un des objectifs de cet article sera de présenter dans un formalisme commun une synthèse de certaines stratégies, et de les comparer expérimentalement.

L'article suit le plan suivant. La section 2 présente les différentes approches permettant d'extraire le flot à partir du TS, à commencer par l'approche standard de Lucas et Kanade [19]. La raison de commencer par la phase de post-traitement et [19] est que cela permet de montrer comment les TS apparaissent de manière naturelle dans l'estimation du flot. Deux autres approches sont ensuite présentées, l'une basée sur l'étude des éléments propres, l'autre sur une formulation voisine de [19] utilisant le flot normal et des mesures de cohérence. La section 3 porte sur les processus de diffusion par EDP non linéaire pour l'obtention d'un champ tensoriel lissé. Nous y présentons aussi une nouvelle approche dans laquelle le coefficient de diffusion dépend d'une mesure de coin. Enfin, la section 4 présente un certain nombre de résultats expérimentaux pour donner une idée des performances ce type d'approches.

2 Flot optique et tenseur de structure

Nous noterons D le domaine spatio-temporel correspondant à la séquence $u(x_1, x_2, t) : D \rightarrow \mathbb{R}$, Ω le domaine spatial des (x_1, x_2) et $\sigma(x_1, x_2, t) : D \rightarrow \mathbb{R}^2$ le champ de vitesse instantané au temps t . Etant donné u le problème est de trouver σ au temps t . Une hypothèse classique consiste à supposer l'intensité constante le long des trajectoires, soit $u(x_1(t), x_2(t), t) = cte$, ce que l'on peut dériver formellement pour obtenir la Contrainte du Flot Optique (CFO) :

$$\nabla u \cdot \sigma + u_t = \nabla_{xt} u \cdot \tilde{\sigma} = 0, \quad (1)$$

où $\tilde{\sigma} = (\sigma_1, \sigma_2, 1)^t$. Cette équation scalaire ne permet que d'obtenir la composante $\sigma_\perp$, normale aux isophotes, i.e. parallèle à $\hat{n} = \nabla u / \|\nabla u\|$, c'est le *problème d'ouverture*. De nombreuses approches ont été proposées pour incorporer une connaissance *a priori* sur la solution, comme par exemple la prise en compte de critères de régularité sur la solution ou de modèles, pour résoudre le problème d'ouverture et restreindre le champ de σ admissibles.

2.1 L'approche de Lucas et Kanade [19]

La solution proposée par Lucas et Kanade [19] est de faire l'hypothèse que le flot est localement constant sur un voisinage et de chercher quel flot vérifie au mieux les équations de CFO dans ce voisinage. Plus précisément, la méthode consiste à résoudre la CFO dans le sens des moindres carrés dans une fenêtre $W(x) \subset \Omega$: à une position x donnée, le flot σ est obtenu comme minimum de l'énergie :

$$\int_{W(x)} (\nabla_{xt} u \cdot \tilde{\sigma})^2 dx' = \int_{W(x)} \tilde{\sigma}^t (\nabla_{xt} u) (\nabla_{xt} u)^t \tilde{\sigma} dx'. \quad (2)$$

Au lieu d'avoir un simple moyennage sur un voisinage x , [19] propose de pondérer la CFO par une fonction $h(x)$ qui

est habituellement un noyau gaussien de moyenne nulle et d'écart type s_x :

$$h(x) = \frac{1}{2\pi s_x^2} e^{-\frac{x^T x}{2s_x^2}}.$$

Si l'on note

$$\langle f \rangle(x, t) = \int_{W(x)} h(x - x') f(x', t) dx' \quad \text{et} \quad (3)$$

$$S = \langle (\nabla_{xt} u) (\nabla_{xt} u)^t \rangle, \quad (4)$$

qui est par définition le Tenseur de Structure (TS), alors le problème (2) peut se réécrire ($\tilde{\sigma}$ est constant sur Ω) :

$$\text{Arg min}_{\tilde{\sigma}=(\sigma_1, \sigma_2, 1)^T} \tilde{\sigma}^T S \tilde{\sigma}. \quad (5)$$

La minimisation de (5) est un problème de moindres carrés pondérés et le résoudre revient donc à considérer le système d'équations :

$$A \begin{bmatrix} \sigma_1 \\ \sigma_2 \end{bmatrix} = b \quad \text{où} \quad (6)$$

$$A = \begin{bmatrix} \langle u_{x_1}^2 \rangle & \langle u_{x_1} u_{x_2} \rangle \\ \langle u_{x_1} u_{x_2} \rangle & \langle u_{x_2}^2 \rangle \end{bmatrix} \quad \text{et} \quad b = \begin{bmatrix} \langle u_{x_1} u_t \rangle \\ \langle u_{x_2} u_t \rangle \end{bmatrix}.$$

Moyenner (1) sur une petite fenêtre peut être insuffisant si le problème d'ouverture persiste, c'est-à-dire que la structure est linéaire et donc $\det(A) = 0$. Une solution envisageable est d'augmenter la régularisation locale en remplaçant dans (2) le lissage spatial par un lissage spatio-temporel. Cela conduit au même système (6) mais où $\langle f \rangle$ est remplacé par $\langle\!\langle f \rangle\!\rangle$ défini par :

$$\langle\!\langle f \rangle\!\rangle(x, t) = \int_{W(x, t)} h(x - x', t - t') f(x', t') dx' dt \quad (7)$$

où $h(x, t)$ est un noyau de lissage spatio-temporel du type :

$$h(x, t) = \frac{1}{2\pi s_x^2} \frac{1}{\sqrt{2\pi s_t^2}} e^{-\frac{x^T x}{2s_x^2} - \frac{t^2}{2s_t^2}}$$

s_t désignant l'écart type temporel. Cette méthode correspond aux travaux de [5].

2.2 L'analyse des éléments propres

Au lieu de résoudre le problème de minimisation (5), on peut chercher la solution du problème voisin

$$\text{Arg min}_{\mu=(\mu_1, \mu_2, \mu_3)^T} \mu^T S \mu \quad (8)$$

où la contrainte $\mu^T \mu = 1$ est ajoutée pour éviter la solution triviale $\mu = 0$ (remarquez que S sera être en général calculé par un lissage spatio-temporel (7)). Le flot correspondant sera donné par :

$$\sigma = \mu_3^{-1} \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}.$$

En incorporant la contrainte à l'aide d'un multiplicateur de Lagrange λ , on se ramène à résoudre le problème de minimisation :

$$\text{Arg min}_{\mu=(\mu_1, \mu_2, \mu_3)^T} L(\mu, \lambda) = \mu^T S \mu + \lambda(1 - \mu^T \mu). \quad (9)$$

La différenciation de (9) par rapport à μ conduit à la condition d'optimalité :

$$S\mu = \lambda\mu \quad (10)$$

i.e. λ doit être une valeur propre de S et μ son vecteur propre correspondant. Pour un tel couple (μ, λ) on a $L(\mu, \lambda) = \lambda$ et donc le minimum de (8) est atteint quand λ est la valeur propre la plus faible de S . Puisque S est le résultat d'une intégration de tenseurs 3×3 positifs semi-définis $T(x, t) = \nabla_{xu} \nabla_{xu}^T$, S est de même nature et possède donc trois valeurs propres réelles positives $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$.

La formulation (8)-(10) est un problème de type moindres carrés totaux pondérés et ce type d'approche pour le flot a été notamment étudié dans [17, 15].

Analyse qualitative. On peut donc analyser le problème du mouvement de la façon suivante, en considérant quatre cas [15] :

- $\lambda_1 = \lambda_2 = \lambda_3 = 0$. Dans cette situation, il n'y a pas de variation en niveaux de gris dans le voisinage (x, t) et on ne peut donc pas distinguer de mouvement apparent.
- $\lambda_1 \gg 0, \lambda_2 = \lambda_3 = 0$. Cela signifie qu'il y a une famille de solution 1D et que le problème d'ouverture n'est pas résolu. L'information contenue dans le voisinage du point (x, t) n'a pas été suffisante pour supprimer l'ambiguïté mais on peut extraire un flot normal :

$$\sigma_{\perp} = \frac{\mu_{1,3}}{\mu_{1,1}^2 + \mu_{1,2}^2} \begin{bmatrix} \mu_{1,1} \\ \mu_{1,2} \end{bmatrix} \quad (11)$$

où μ_1 est le vecteur propre correspondant à la plus grande valeur propre λ_1 .

- $\lambda_1 \gg 0, \lambda_2 \gg 0, \lambda_3 = 0$ indique que le voisinage de (x, t) contenait suffisamment d'informations pour donner une direction unique et résoudre le problème d'ouverture. Le flot est donné par :

$$\sigma = \mu_{3,3}^{-1} \begin{bmatrix} \mu_{3,1} \\ \mu_{3,2} \end{bmatrix}$$

où μ_3 est le vecteur propre correspondant à la valeur propre unique $\lambda_3 = 0$.

- $\lambda_1 \gg 0, \lambda_2 \gg 0, \lambda_3 \gg 0$ correspond au cas où aucun mouvement cohérent ne peut être déterminé ce qui correspond à une discontinuité du flot.

Cette approche reste théorique et une implémentation a été par exemple proposée par Nagel et Middendorf [21, 20]. Ayant choisis des seuils τ_n et τ_g , ils définissent un *mouvement neutre* si $\lambda_1 + \lambda_2 + \lambda_3 \leq \tau_n$. Le cas de *direction dominante du gradient* correspond au problème d'ouverture où $\frac{\lambda_2 + \lambda_3}{\lambda_1 + \lambda_2 + \lambda_3} \leq \tau_g$. Néanmoins, si $\frac{\lambda_2}{\lambda_3} \geq 50$, le mouvement

peut être estimé de façon fiable. Le cas $\lambda_3 \gg 0$ correspond à une *discontinuité du mouvement*, c'est-à-dire une zone de rencontre entre deux régions subissant des mouvements différents. On retrouve dans cette classification les cas précédents. L'intérêt de leur approche réside dans la mise en œuvre de seuils pour segmenter le flot en pratique.

Interprétation géométrique. La figure 1 permet d'illustrer les diverses configurations possibles lors de l'évaluation du flot. Le premier cas correspond à la situation idéale. On peut en effet voir l'approche par moindres carrés totaux comme l'estimation des 3 paramètres décrivant un plan de $\mathbb{R}^3$. Si, dans le voisinage considéré, les vecteurs propres sont suffisamment coplanaires, le flot calculé est fiable. Si par contre les vecteurs propres sont principalement orientés dans une direction (celle du gradient), une famille de solution existe, c'est le problème d'ouverture. Enfin, si les vecteurs ne présentent aucune cohérence, ils décrivent un domaine isotrope. Ceci correspond à des régions trop homogènes ou d'occultations (en fonction de la taille des tenseurs).

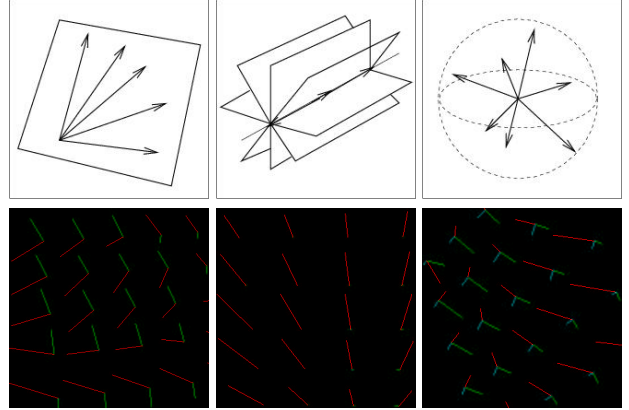


FIG. 1 – Répartition des vitesses dans un voisinage spatial dans les cas coplanaire, anisotrope et isotrope, et exemples de champs tensoriels réels dans ces configurations (seuls les axes principaux sont affichés, $\lambda_1 v_1$ en rouge, $\lambda_2 v_2$ en vert et $\lambda_3 v_3$ en cyan).

Définition des mesures de cohérences. L'étude de la section 2.2 a montré comment utiliser les valeurs propres pour identifier la nature du problème de flot optique. Pour mieux quantifier les valeurs relatives des valeurs propres dans les différentes situations, il est d'usage d'introduire les quantités suivantes, comprises entre 0 et 1 :

- La *mesure de cohérence totale* définie par

$$c_t = \left(\frac{\lambda_1 - \lambda_3}{\lambda_1 + \lambda_3} \right)^2. \quad (12)$$

Elle représente la confiance en l'estimation du mouvement. Quand c_t est proche de 1, on peut alors estimer un mouvement qui correspondra soit au flot normal soit au flot optique. Si c_t est proche de 0 cela indique qu'il n'y a pas de mouvement cohérent dans le voisinage de (x, t) .

- La mesure de cohérence spatiale définie par

$$c_s = \left(\frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2} \right)^2 \quad (13)$$

symbolise la présence du problème d'ouverture. Une valeur de c_s proche de 1 signifie que le problème d'ouverture est présent. Dans toutes les autres situations, c_s sera faible.

- La mesure de coin c_c définie par

$$c_c = c_t - c_s. \quad (14)$$

Quand c_c est proche de 1, il n'y pas de problème d'ouverture et le flot peut être déterminé sans ambiguïté. Cette situation correspond au cas coplanaire de la figure 1. Dans la suite de cet article nous proposerons une utilisation directe de c_c dans une processus de diffusion tensorielle.

2.3 Utilisation du flot normal et des mesures de cohérences

La présence d'un problème d'ouverture permet quand même le calcul du flot normal $\sigma_\perp$ en un point donné, par la formule (11) et il est lié au flot optique par :

$$\sigma_\perp = \sigma^T \hat{n} \quad (15)$$

où $\hat{n} = \nabla u / \|\nabla u\|$ est la direction du gradient spatial normalisé qui peut aussi être calculée à partir du tenseur de structure par

$$\hat{n} = \begin{bmatrix} n_1 \\ n_2 \end{bmatrix} = \frac{1}{\sqrt{\mu_{1,1}^2 + \mu_{1,2}^2}} \begin{bmatrix} \mu_{1,1} \\ \mu_{1,2} \end{bmatrix} \quad (16)$$

où μ_1 est le vecteur propre correspondant à la plus grande valeur propre de S . Alors, en considérant un voisinage un peu plus grand $W(x, t)' \supseteq W(x, t)$ de (x, t) , avec un nouveau noyau de lissage $h'(x, t)$, on peut calculer σ avec (15) dans le sens des moindres carrés, comme suit :

$$\sigma = \underset{s=(s_1, s_2)^T}{\text{Argmin}} \langle c \rangle^{-1} \int_{W'(x, t)} h'(x'-x, t'-t) c(x') [s^T \hat{n}(x) - \sigma_\perp(x)]^2 dx' dt' \quad (17)$$

où $c(x')$ est une mesure de confiance en l'estimation du flot normal basée sur les mesures de cohérences de la section précédente et ici

$$\langle f \rangle = \int_{W'(x, t)} h'(x'-x, t'-t) f(x', t') dx' dt'.$$

La minimisation est menée en résolvant le système d'équations

$$A \begin{bmatrix} \sigma_1 \\ \sigma_2 \end{bmatrix} = b \quad \text{où} \quad (18)$$

$$A = \begin{bmatrix} \langle cn_1^2 \rangle & \langle cn_1 n_2 \rangle \\ \langle cn_1 n_2 \rangle & \langle cn_2^2 \rangle \end{bmatrix} \quad \text{et} \quad b = \begin{bmatrix} \langle cn_1 \sigma_\perp \rangle \\ \langle cn_2 \sigma_\perp \rangle \end{bmatrix}.$$

Si $\det(A) = \langle cn_1^2 \rangle \langle cn_2^2 \rangle - \langle cn_1 n_2 \rangle^2 = 0$, le problème d'ouverture persiste dans $W'(x, t)$, les vecteurs normaux dans $W'(x, t)$ sont toujours collinéaires.

Dans cette démarche, citons aussi les travaux [12, 24, 30] qui sont basés sur des idées voisines et proposent différentes mesures de confiances ou des modèles polynomiaux d'ordres supérieurs.

2.4 Commentaire sur ces techniques

Dans la section 2, trois méthodes d'extraction de flot en un point (x, t) ont été introduites.

- La première est une simple estimation en moindres carrés pondérés pour le système formé des contraintes de flot optique au voisinage de (x, t) . Cela correspond à une hypothèse implicite de grande précision sur le calcul des dérivées spatiales de l'image, l'erreur étant supposée porter uniquement sur les dérivées temporelles. C'est la la plus simple à mettre en œuvre et est *numériquement la plus stable*. L'hypothèse implicite de précision sur les dérivées spatiales est néanmoins rarement vérifiée sur des exemples réels, et ceci entraîne un biais sur le résultat : le flot estimé a tendance à être de norme inférieure au flot réel et biaisé dans la direction du flot normal. De plus, dans le cas bruité, elle fournit en général un flot même si celui-ci n'est pas défini.
- La seconde méthode est quant à elle une estimation en moindres carrés totaux pondérés, sur le système formé des différentes contraintes de flot optique au voisinage de (x, t) , une erreur de mesure sur les dérivées spatiales étant cette fois-ci prise en compte. Cette approche est aussi en général biaisée. Néanmoins, lorsque le bruit est homogène, une estimation de l'orientation non biaisée peut être obtenue, ce qui n'est pas le cas pour la norme, qui aura aussi tendance à être sous-estimée. Dans la pratique, nous avons rencontré quelques problèmes de stabilité numérique.
- La troisième méthode présuppose l'extraction du flot normal au voisinage de (x, t) et estime le flot optique par une méthode de moindres carrés pondérés pour le système de projection du flot optique sur le flot normal. Cette méthode cherche à prendre en compte l'erreur dans l'estimation locale du flot normal. Comme les méthodes précédentes, elle s'avère biaisée en présence de bruit mais peut fournir des estimations plus précises que la deuxième méthode. La difficulté de cette approche réside notamment dans le choix du coefficient de confiance sur le flot normal.

Ces problèmes d'estimations sont rencontrés dans de nombreuses situations [28] et pour le cas du flot optique le lecteur pourra consulter [16, 6, 11].

3 Méthodes de diffusion non linéaire pour trouver la structure

3.1 Motivation

Dans la section précédente, nous avons étudié trois modèles locaux pour extraire le flot optique. Leur principe est d'estimer la vitesse en intégrant une équation de type CFO sur un voisinage, la contrainte étant en général pondérée par une gaussienne. Cela revient à lisser le champ des tenseurs associé, ce qui peut être fait dans le domaine spatial (3) ou spatio-temporel (7). Ainsi, minimiser ces énergies conduit à des systèmes algébriques où les coefficients moyennés des TS sont utilisés.

Si l'on envisage le lissage gaussien comme un processus de diffusion linéaire, l'idée est alors de considérer des méthodes non linéaires pour obtenir des coefficients des TS lissés, ce que nous allons étudier dans cette section. L'idée sera alors d'utiliser les coefficients de ces tenseurs lissés à la place de ceux obtenus par convolution gaussienne (cf section 2).

3.2 Synthèse des approches existantes

Dans cette section nous revisitons quelques contributions majeures sur des traitement de champs de tenseurs pour l'estimation du flot optique.

Spies et Scharf [25]. Les auteurs proposent une méthode combinant le lissage des données images et le calcul des tenseurs. L'équation sur la séquence u est gouvernée par le champ de tenseurs dans le but de rehausser les structures spatio-temporelles. Cette méthode entre dans le formalisme des approches proposées par Weickert *et al.* pour l'amélioration d'images [33, 34]. Les auteurs proposent de résoudre :

$$u_t = \text{div}(D \nabla u) \quad (19)$$

où le tenseur D utilise les mêmes vecteurs propres $\{v_i\}$ que ceux du TS (4) :

$$D = \sum_{i=1}^3 \gamma_i v_i^t v_i \quad \text{avec} \quad (20)$$

$$\gamma_i = \begin{cases} 1 & \text{si } \lambda_i \leq \alpha_2 \\ 1 - \exp\left(\frac{-\alpha_1}{(\lambda_i - \alpha_2)^2}\right) & \text{sinon} \end{cases} \quad (21)$$

où $\{\alpha_i\}_{i=1,2} > 0$ sont des constantes (voir [25]). Comparée aux modèles initiaux de Weickert [33, 34], cette approche traite toutes les directions de façon individuelle et se révèle plus efficace aux discontinuités du mouvement. Les TS résultants sont régularisés par un lissage gaussien et le flot est extrait par la méthode standard d'analyse de valeurs propres développée dans la section 2.2.

Brox et Weickert [7]. Les auteurs présentent une méthode de diffusion non linéaire sur les matrices qui a été appliquée au TS. L'équation est :

$$(S_{ij})_t = \text{div}(D \nabla S_{ij}) \quad (22)$$

où le tenseur de diffusion D est défini par :

$$D = \sum_{i=1}^3 \gamma_i w_i^t w_i \quad \text{avec} \quad (23)$$

$$\gamma_i = h(\eta_i), \quad h(s) = 1 - \exp\left(-\frac{\alpha}{s^4}\right), \quad (24)$$

où α est une constante fixée selon [31] et (η_i, w_i) sont les éléments propres de N défini par :

$$N = (\nabla^\sigma n)(\nabla^\sigma n)^t \quad \text{avec } n = \sqrt[4]{\sum_{kl} s_{kl}^2}. \quad (25)$$

La fonction n est la racine de la norme de Frobenius du tenseur et au temps initial, elle correspond à la norme du gradient de l'image. Cela représente une façon indirecte de coupler le lissage des TS avec la structure de l'image. D'un point de vue numérique, après avoir stoppé la diffusion à un temps t qui constitue un paramètre déterminant, le flot est extrait à partir de l'approche décrite dans la section 2.1 où les coefficients de (6) ont été remplacés par ceux des TS lissés de manière non linéaire.

Nagel *et al.* [22, 21, 20]. Cette approche repose sur le fait qu'un noyau gaussien multidimensionnel constitue un opérateur de voisinage linéaire dont l'orientation et l'étendue spatiale peuvent être adaptées à la structure locale de l'image. L'aspect adaptatif de cette méthode se traduit par l'utilisation du tenseur de structure pour raffiner l'orientation et l'étendue spatio-temporelle d'un noyau gaussien à une position (x, t) . Pratiquement, le tenseur de structure T est calculé itérativement par :

$$T(x, t) = ((\nabla u)(\nabla u)^T \star G_{\Sigma_M})(x, t)$$

avec

$$\nabla u(x, t) = (u \star \nabla G_{\Sigma_G})(x, t)$$

Σ_G étant un pseudo-inverse du tenseur de structure obtenu à l'itération précédente et $\Sigma_M \approx 2\Sigma_G$. Afin d'éviter les solutions dégénérées, les valeurs propres subissent une transformation qui les projette dans un intervalle dépendant essentiellement du nombre d'images dans la séquence et de seuils minimaux et maximaux. Elle permet d'évaluer la qualité du flot grâce aux critères qualitatifs proposés par les auteurs et rappelés dans la section 2.2. L'inconvénient majeur de cette méthode réside dans ses temps de calcul extrêmement élevés et ses besoins conséquents en mémoire. Le calcul exhaustif des masques de convolution orientés non-séparables en est la cause principale.

Cette méthode basée sur du filtrage gaussien présente des similarités avec [25] et [7]. Elle cherche à estimer de manière couplée l'image et un champ de tenseurs lissés. L'image est lissée par un pseudo-inverse du tenseur de structure comparable à (20). Le TS est à son tour lissé en utilisant ce pseudo-inverse selon un processus de diffusion comparable au formalisme EDP présenté dans [35].

Méthodes pour l'IRMd. L'amélioration des tenseurs de structure intervient dans d'autres applications comme en IRM de diffusion où l'on s'intéresse à l'extraction des fibres dans la matière blanche correspondant aux lignes de champ des directions principales des tenseurs. [9, 10, 26, 27, 29]. A cause de leur utilisation *a posteriori*, les opérations effectuées sur ces tenseurs ont pour but de rehausser des structures unidimensionnelles dans des directions quelconques, ce qui ne correspond pas à notre objectif, où l'on cherche à obtenir des tenseurs plateaux (cf figure 1).

3.3 Vers une utilisation des mesures de cohérence

Dans la section 2.2, plusieurs quantités permettant de savoir si un flot peut être extrait de manière fiable à partir des TS. En particulier, il paraît important d'avoir une mesure de coin élevée pour pouvoir résoudre le problème d'ouverture. Une solution possible pour augmenter la cohérence est d'utiliser des gaussiennes de variance croissantes afin de récupérer de la structure. Pour l'illustrer, nous montrons dans la figure 2 l'évolution de la mesure de coin d'un champ de tenseur associé à une séquence en fonction de la variance des gaussiennes. La moyenne commence par augmenter ce qui signifie que l'on récupère de la structure, avant de diminuer car trop d'information est prise en compte. Cette évolution montre donc qu'il est important de savoir stopper la diffusion dès que suffisamment d'information a été incorporée dans le TS. D'autre part, il semble clair que la taille du voisinage que l'on souhaite prendre en compte doit dépendre de la configuration locale de la scène. Ces remarques peuvent être transcrites en terme d'EDP et l'on propose de diffuser le champ de TS en utilisant la cohérence, selon :

$$(S_{ij})_t = \text{div}(g(c_c)\nabla S_{ij}) \quad (26)$$

où g est une fonction continue telle que :

$$g_\alpha(s) = \begin{cases} 1 & \text{si } s \leq \alpha - \varepsilon \\ (s - \alpha + \varepsilon)/(2\varepsilon) & \text{si } \alpha - \varepsilon \leq s \leq \alpha + \varepsilon \\ 0 & \text{si } s \geq \alpha + \varepsilon \end{cases} \quad (27)$$

où $\varepsilon = 0.1$ et $\alpha = 0.9$ dans les expériences. Cette approche rentre dans le cadre des méthodes proposées par Brox et Weickert qu'ils résument dans [35]. Notre contribution vient ici de l'introduction de la mesure de coin de manière simple et explicite dans le processus de diffusion. L'interprétation de (26)-(27) est qu'une diffusion isotrope va se poursuivre jusqu'à ce que de la structure permettant de résoudre la CFO soit trouvée. La différence de comportement de l'évolution de la cohérence est illustrée dans la figure 2.

4 Résultats expérimentaux

Protocole. Les méthodes que nous avons cherchées à comparer sont :

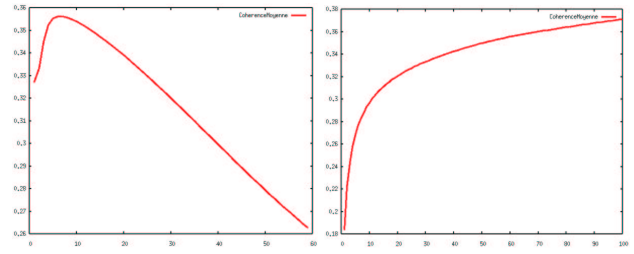


FIG. 2 – Evolution de la cohérence pour un champ d'un champ de tenseur lissé avec des gaussiennes de taille croissante (Gauche) ou au cours des itérations de l'approche proposée.

- $[G_x G_x]$ la méthode qui consiste en un lissage spatial par une gaussienne du volume des images et des tenseurs,
- $[G_{xt} G_{xt}]$ la même méthode que précédemment mais avec un lissage spatio-temporel,
- [25] la méthode de Spies et Scharf,
- [7] la méthode de Brox et Weickert,
- [22, 21, 20] la méthode de Nagel *et al.*,
- [DTMC] la méthode de diffusion proposée (26).

Toutes les séquences que nous avons utilisées sont des séquences synthétiques pour lequel le mouvement réel est connu. Cela permet donc de faire des comparaisons entre le flot réel et le flot calculé. Les méthodes de type TS étant peu adaptées aux méthodes multiéchelles, nous avons choisi des séquences avec de faibles déplacements. Les expériences effectuées semblent confirmer un meilleur comportement des méthodes de type TS par rapport aux méthodes globales. Cela n'est plus vrai pour des séquences qui comportent des mouvements importants.

Les critères d'erreur que nous avons utilisés sont des critères d'erreur standard dans la littérature, l'erreur angulaire, calculée de la même façon que dans [4] :

$$\Psi_e = \arccos \left(\frac{u_r u_c + v_r v_c + 1}{\sqrt{(u_r^2 + v_r^2 + 1)(u_c^2 + v_c^2 + 1)}} \right)$$

et l'erreur en norme

$$N_e = \sqrt{(u_r - u_c)^2 + (v_r - v_c)^2}$$

où (u_r, v_r) est le flot réel et (u_c, v_c) le flot calculé.

Résultats. Nous commençons par présenter des calculs d'erreurs sur deux séquences test synthétiques. Les résultats sont donnés dans les figures 3, 4 et 5 et la table 1 :

- Dans la séquence *Street* les erreurs sont d'ordre équivalent alors que la séquence *Yosemite* révèle des différences plus importantes. L'approche proposée [DTMC] donne les meilleurs résultats. Une possible explication est la présence de beaucoup structure non linéaire (i.e. de coins) favorisant une diffusion basée sur c_c .
- Afin de mieux situer les performances des approches type TS par rapport aux méthodes de minimisation cherchant directement un flot régularisé, nous avons testé

les approches [13, 1, 2, 23]. Les meilleurs résultats sont obtenus avec les méthodes type TS. On observe cependant une différence de comportement en fonction du cas traité. Les résultats inférieurs de la plupart des approches [13, 1, 2, 23] dans le cas de la séquence *Street* sont essentiellement dus à une propagation excessive de l'erreur dans la direction du mouvement (cf figure 3).

(Notez que les résultats obtenus pour la séquence *Street* sont moins bons que ceux présentés dans [7] car nous avons utilisé une sous-séquence, autour du véhicule en mouvement).

La figure 6 représente l'orientation des flots pour une séquence synthétique où deux textures évoluent en sens opposé. L'idée est d'estimer qualitativement la localisation des discontinuités. Dans cette optique, certaines approches globales comme [2] permettent une bonne localisation des discontinuités (voir aussi dans la figure 4). Les erreurs d'orientations au voisinage des discontinuités sont présentes dans toutes les méthodes de type TS et de manière importante pour [22, 21, 20], pour des raisons de stabilité des moindres carrés totaux (il est d'ailleurs à noter que ces méthodes cherchent plus à *segmenter* le champ de mouvement qu'à l'estimer à proprement parler).

La figure 7 illustre pour l'approche [DTMC] l'évolution de la cohérence et de l'erreur en fonction des itérations. On peut observer la disparition progressive des erreurs et la conservation de mesures faibles dans les régions les moins texturées. Expérimentalement, on observe également une augmentation de la cohérence moyenne.

5 Conclusion

Dans cet article, nous avons cherché à regrouper sous un même formalisme certaines des approches récemment proposées dans la littérature. Cela a permis de révéler un certain nombre de relations, notamment concernant les méthodes d'obtention de tenseurs de structure lissés. Dans ce cadre nous avons proposé une approche simple de diffusion qui prend en compte directement la mesure de coin. Cette présentation des différentes approches a été complétée par un jeu d'expérimentations visant à les comparer quantitativement. Ces expériences illustrent le potentiel des méthodes du type TS et de celle que nous avons proposées et constituent une base expérimentale que nous souhaitons développer. Notre objectif sera de poursuivre de manière plus systématique cette étude et tenter d'analyser les méthodes de diffusion en IRMd qui manipulent les mêmes objets avec des objectifs différents.

Remerciements

Nous tenons à remercier les auteurs de [7] et [1] pour leur aide dans la génération des résultats obtenus avec leur approche.

Street			Yosemite		
Méthode	Ψ_e	N_e	Méthode	Ψ_e	N_e
<i>Méthodes basées sur le TS</i>					
$[G_x G_x]$	10.7	0.32	$[G_x G_x]$	4.90	0.28
$[G_{xt} G_{xt}]$	10.0	0.35	$[G_{xt} G_{xt}]$	4.36	0.26
[25]	10.52	0.29	[25]	16.2	0.77
[7]	9.4	0.27	[7]	8.87	0.57
[DTMC]	10.2	0.30	[DTMC]	2.82	0.16
[21, 20]	16.5	0.5	[21, 20]	13.65	1.03
<i>Méthodes variationnelles globales</i>					
[13]	33.6	0.79	[13]	16.4	0.82
[1]	18.6	0.43	[1]	4.0	0.26
[2]	28.3	0.69	[2]	15.8	0.77
[23]	20.7	0.51	[23]	7.9	0.49

TAB. 1 – Calculs d'erreurs sur séquences synthétiques

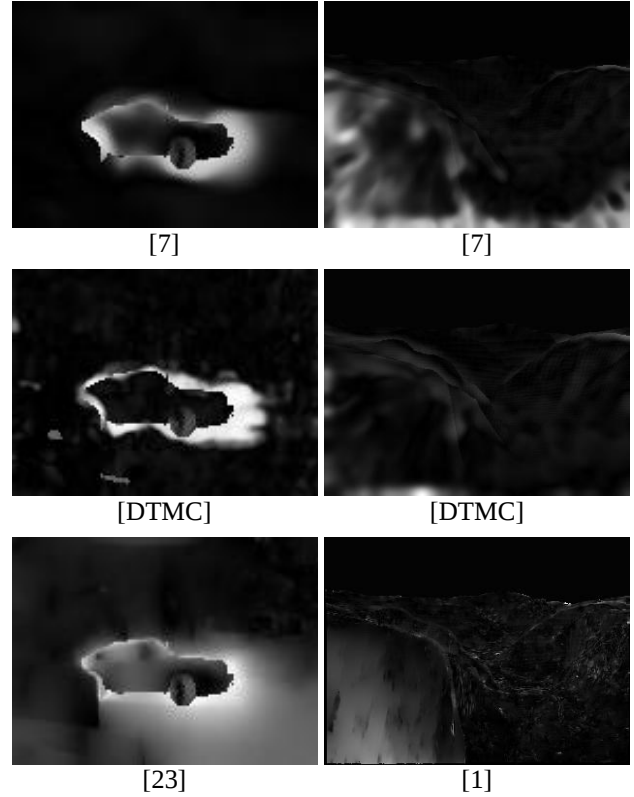
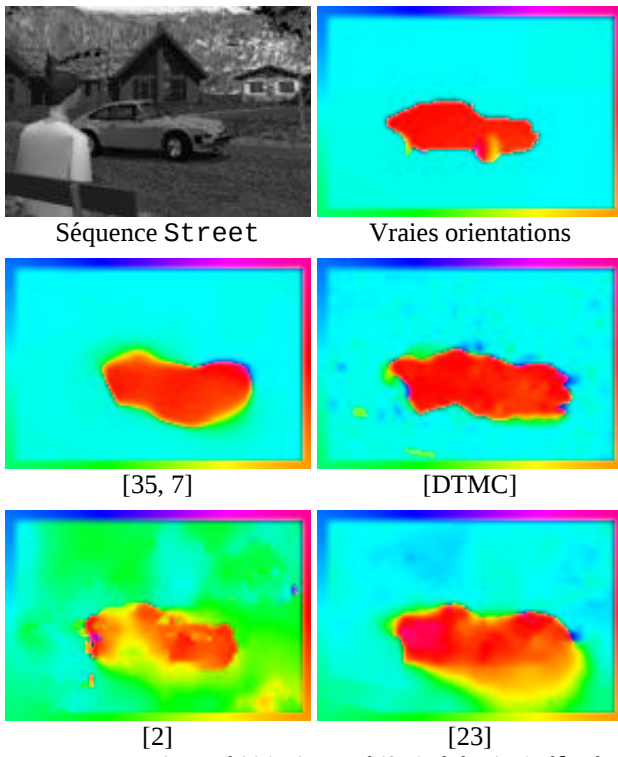


FIG. 3 – Images d'erreurs angulaires correspondant aux mesures du tableau 1



www.cs.otago.ac.nz/research/vision/Research/OpticalFlow/opticalflow.html

FIG. 4 – Séquence synthétique avec un mouvement occulté de la voiture sur un fond mobile dans la direction opposée.

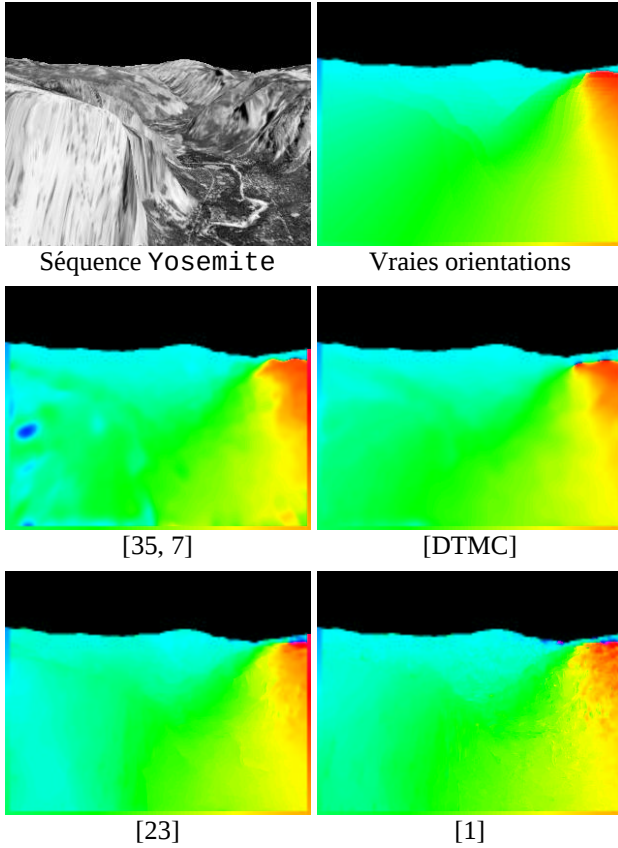


FIG. 5 – Séquence Yosemite où les nuages n'ont pas été considérés dans les calculs.

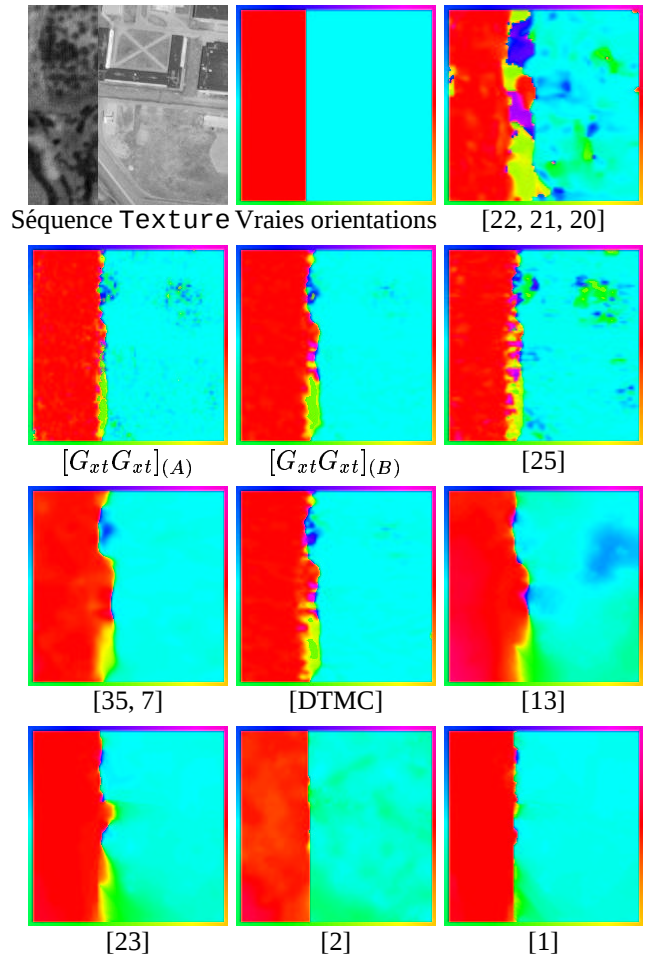


FIG. 6 – Estimation des orientations du flot sur une séquence synthétique où deux textures se déplacent en sens inverse. Pour les résultats obtenus avec $[G_{xt}G_{xt}]$ les valeurs des variances des noyaux gaussien est (A) :pré=(1,0.5), post=(2,1), (B) :pré=(1,0.5), post=(1,0.5).

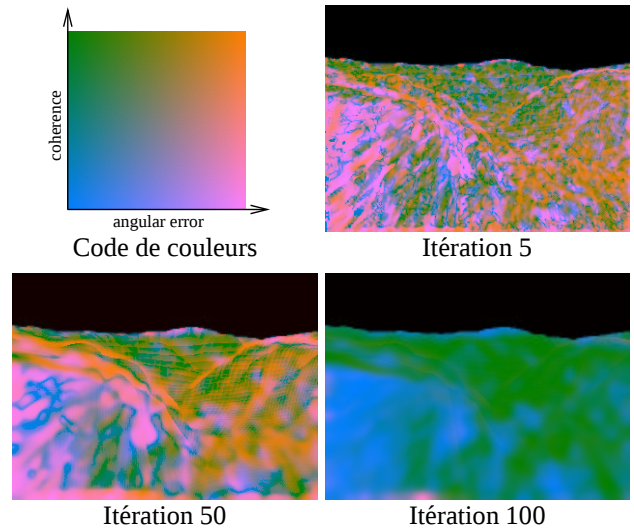


FIG. 7 – Evolution de l'erreur et de la cohérence sur la séquence Yosemite avec la méthode [DTMC]

Références

- [1] L. Alvarez, R. Deriche, T. Papadopoulos, and J. Sanchez. Symmetrical dense optical flow estimation with occlusion detection. In A. Heyden, G. Sparr, M. Nielsen, and P. Johansen, editors, *Proceedings of the 7th European Conference on Computer Vision*, pages 721–735, Copenhagen, Denmark, May 2002. Springer-Verlag.
- [2] G. Aubert, R. Deriche, and P. Kornprobst. Computing optical flow via variational techniques. *SIAM Journal of Applied Mathematics*, 60(1) :156–182, 1999.
- [3] G. Aubert and P. Kornprobst. *Mathematical Problems in Image Processing : Partial Differential Equations and the Calculus of Variations*, volume 147 of *Applied Mathematical Sciences*. Springer-Verlag, January 2002.
- [4] J.L. Barron, D.J. Fleet, and S.S. Beauchemin. Performance of optical flow techniques. *The International Journal of Computer Vision*, 12(1) :43–77, 1994.
- [5] J. Bigun, G. H. Granlund, and J. Wiklund. Multi-dimensional orientation estimation with applications to texture analysis and optical flow. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(8) :775–790, August 1991. Report LiTH-ISY-I-0828 1986 and Report LiTH-ISY-I-1148 1990, both at Computer Vision Laboratory, Linköping University, Sweden.
- [6] Jacques Bride and Peter Meer. Global registration via direct methods : a statistical approach. In *Proceedings of CVPR'01*, December 2001.
- [7] T. Brox and J. Weickert. Nonlinear matrix diffusion for optic flow estimation. In *DAGM-Symposium*, pages 446–453, 2002.
- [8] Hans Burkhardt and Bernd Neumann, editors. volume 1406–1407 of *Lecture Notes in Computer Science*, Freiburg, Germany, June 1998. Springer-Verlag.
- [9] O. Coulon, D.C. Alexander, and S.R. Arridge. A geometrical approach to 3d diffusion tensor magnetic resonance image regularisation. Technical report, Department of Computer Science, University College London., 2001.
- [10] O. Coulon, D.C. Alexander, and S.R. Arridge. A regularization scheme for diffusion tensor magnetic resonance images. In *XVIIth International Conference on Information Processing in Medical Imaging*, 2001.
- [11] C. Fermüller, D. Shulman, and Y. Aloimonos. The statistics of optical flow. *Computer Vision and Image Understanding*, 82(1) :1–32, April 2001.
- [12] L. Florack, W. Niessen, and M. Nielsen. The intrinsic structure of optic flow incorporating measurement duality. *ijcv*, 27(3) :263–286, 1998.
- [13] B.K. Horn and B.G. Schunk. Determining optical flow. *Artificial Intelligence*, 17 :185–203, 1981.
- [14] IEEE Computer Society. Vancouver, Canada, 2001. IEEE Computer Society Press.
- [15] B. Jähne. *Handbook of computer vision*, vol. 2 : *signal processing and pattern recognition*. Academic press, 1999.
- [16] B. Jähne and H. Haußecker. Performance characteristics of low-level motion estimators in spatiotemporal images. In *Performance Characterization in Computer Vision*, 9. *Theoretical Foundations of Computer Vision, TFCV, March 1998, Dagstuhl, Germany*. Kluwer, 1998.
- [17] B. Jähne, H. W. Haussecker, H. Scharr, H. Spies, D. Schmudt, and U. Schurr. Study of dynamical processes with tensor-based spatiotemporal image processing techniques. In Burkhardt and Neumann [8], pages 322–336.
- [18] G. Kühne, J. Weickert, O. Schuster, and S. Richter. A tensor-driven active contour model for moving object segmentation. In *Proceedings of the International Conference on Image Processing*, volume 2. IEEE Signal Processing Society, 2001.
- [19] B. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *International Joint Conference on Artificial Intelligence*, pages 674–679, 1981.
- [20] M. Middendorf and H.-H. Nagel. Estimation and interpretation of discontinuities in optical flow fields. In *Proceedings of the 8th International Conference on Computer Vision* [14], pages 178–183.
- [21] M. Middendorf and H.-H. Nagel. Empirically convergent adaptive estimation of grayvalue structure tensors. In *DAGM-Symposium*, pages 66–74, 2002.
- [22] H.-H. Nagel and A. Gehrke. Spatiotemporally adaptive estimation and segmentation of OF-fields. In Burkhardt and Neumann [8], pages 86–102.
- [23] H.H. Nagel and W. Enkelmann. An investigation of smoothness constraint for the estimation of displacement vector fields from images sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8 :565–593, 1986.
- [24] K. S. Pedersen and M. Nielsen. Computing optic flow by scale-space integration of normal flow. In Michael Kerckhove, editor, *Scale-Space and Morphology in Computer Vision*, pages 14–25. Springer-Verlag, July 2001.
- [25] H. Spies and H. Scharr. Accurate optical flow in noisy image sequences. In *Proceedings of the 8th International Conference on Computer Vision* [14], pages 587–592.
- [26] D. Tschumperlé and R. Deriche. Diffusion tensor regularization with constraints preservation. In *IEEE*

Computer Society Conference on Computer Vision and Pattern Recognition, Kauai Marriott, Hawaii, December 2001.

- [27] D. Tschumperlé and R. Deriche. Regularization of orthonormal vector sets using coupled PDE's. In *IEEE Workshop on Variational and Level Set Methods*, pages 3–10, Vancouver, Canada, July 2001.
- [28] S. van Huffel and J. Vandewalle. *The Total least squares problem : computational aspects and analysis*, volume 9 of *Frontiers in applied mathematics*. SIAM, 1991.
- [29] B. Vemuri, Y. Chen, M. Rao, T. McGraw, T. Mareci, and Z. Wang. Fiber tract mapping from diffusion tensor mri. In *1st IEEE Workshop on Variational and Level Set Methods in Computer Vision (VLSM'01)*, July 2001.
- [30] H.-Y. Wang and K.-K. Ma. Accurate optical flow estimation using adaptive scale-space and 3d structure tensor. In *Proceedings of the International Conference on Image Processing*, volume 2, pages 301–304. IEEE Signal Processing Society, 2002.
- [31] J. Weickert. *Anisotropic Diffusion in Image Processing*. PhD thesis, University of Kaiserslautern, Germany, Laboratory of Technomathematics, January 1996.
- [32] J. Weickert. A review of nonlinear diffusion filtering. *Scale-Space Theory in Computer Vision, Lecture Notes in Comp. Science (Springer, Berlin)*, 1252 :3–28, 1997. Invited Paper.
- [33] J. Weickert. *Anisotropic Diffusion in Image Processing*. Teubner-Verlag, Stuttgart, 1998.
- [34] J. Weickert. Coherence-enhancing diffusion filtering. *The International Journal of Computer Vision*, 31(2/3) :111–127, April 1999.
- [35] J. Weickert and T. Brox. Diffusion and regularization of vector and matrix-valued images. Technical report, Universitat des Saarlandes, 2002.