Global Optimality of Gradient Vector Flow

Chenyang Xu and Jerry L. Prince[1]
Center for Imaging Science, Department of Elec. and Comp. Eng.
The Johns Hopkins University, Baltimore, MD 21218
e-mail: *chenyang@jhu.edu* and *prince@jhu.edu*

I. INTRODUCTION

In [1, 2], Xu and Prince introduced *gradient vector flow* (GVF), a class of vector fields derived from images, that can be used as external forces for deformable models [3]. Figure 1 illustrates the use of GVF in a deformable model to extract a two-dimensional U-shape object. GVF can be defined through either a variational formulation or a partial differential equation. In this paper, we are concerned with the variational formulation introduced in [2]. The solution to this variational formulation was obtained in [2] by first deriving the necessary condition, the Euler-Lagrange Equation (ELE), and then solving the ELE numerically. Here, we prove the convexity of the GVF variational formulation using the convexity analysis described in [4] and point out that the corresponding ELE is in fact a sufficient condition for *globally* minimizing the variational energy formulation.

II. DEFINITIONS AND BACKGROUND

Let us denote a point in n-dimensional space $\boldsymbol{R}^n$ by $\boldsymbol{x} = (x^1, \cdots, x^n)$, a scalar function at $\boldsymbol{x}$ by $f(\boldsymbol{x}) = f(x^1, \cdots, x^n)$, and a vector function at $\boldsymbol{x}$ by $\boldsymbol{v}(\boldsymbol{x}) = [v^1(x^1, \cdots, x^n), \cdots, v^n(x^1, \cdots, x^n)]^T$. The gradient of $f(\boldsymbol{x})$ yields a vector function that is given by $\nabla f = [\frac{\partial f}{\partial x^1}, \cdots, \frac{\partial f}{\partial x^n}]^T$, and the gradient of $\boldsymbol{v}(\boldsymbol{x})$ yields a tensor that is given by $\nabla \boldsymbol{v} = (\partial v^i / \partial x^j), i, j = 1, \cdots, n$. We denote the Euclidean inner product between vector functions $\boldsymbol{v}$ and $\boldsymbol{u}$ as $\boldsymbol{v} \cdot \boldsymbol{u}$. We also define the inner product between tensors $\boldsymbol{T}$ and $\boldsymbol{S}$ as $\boldsymbol{T} \cdot \boldsymbol{S} = \sum_{i,j=1}^n T_{ij} S_{ij}$. We further assume all these functions are defined in a bounded domain $\Omega \subset \boldsymbol{R}^n$ with $\partial\Omega$ as its boundary.

As described in [2], the n-dimensional GVF is defined as the vector function $\boldsymbol{v}(x)$ in a subset of the Sobolev space denoted by $W_2^2(\Omega)$ [5] that minimizes the following functional

$$J(\boldsymbol{v}) = \int_\Omega g|\nabla \boldsymbol{v}|^2 + h|\boldsymbol{v} - \nabla f|^2 d\boldsymbol{x}, \tag{1}$$

where $g(\boldsymbol{x})$ and $h(\boldsymbol{x})$ are nonnegative functions defined on Ω, $|\nabla \boldsymbol{v}|$ is the vector norm for tensors given by $\sqrt{\nabla \boldsymbol{v} \cdot \nabla \boldsymbol{v}}$, and $|\boldsymbol{v} - \nabla f|$ is the usual Euclidean norm for vectors given by $\sqrt{(\boldsymbol{v} - \nabla f) \cdot (\boldsymbol{v} - \nabla f)}$.

A decisive role in the optimization of a real-valued functional J on a linear space $\mathcal{Y}$ such as the Sobolev space is played by *Gâteaux variations* [4]. Gâteaux variations are directional derivatives of J in $\mathcal{Y}$ that play similar roles as partial derivatives of real-valued function defined on $\boldsymbol{R}^n$.

Definition 1 ([4], p45) For $y, v \in \mathcal{Y}$:

$$\delta J(y; v) \stackrel{\text{def}}{=} \lim_{\varepsilon \to 0} \frac{J(y + \varepsilon v) - J(y)}{\varepsilon}$$

assuming that this limit exists, is called the *Gâteaux variation of J at y in the direction v.*

[1]This work was supported in part by an NSF Presidential Faculty Grant (MIP-9350336) and by an NIH Grant (R01NS37747).

Note that both the definition and the discussion in the rest of the paper is valid for either convex or strictly convex J. For compact notation, the condition for strict convexity is enclosed in brackets.

Definition 2 ([4], p54) A real-valued function J defined on a set $\mathcal{D}(\Omega)$ in a linear space $\mathcal{Y}$ is said to be [strictly] convex on $\mathcal{D}$ provided that when y and $y + v \in \mathcal{D}$ then $\delta J(y; v)$ is defined and $J(y + v) - J(y) \geq \delta J(y; v)$ [with equality iff $v = \mathbf{0}$].

Here **0** stands for the null function.

Proposition 1 ([4], p54) *If J is [strictly] convex on $\mathcal{D}(\Omega)$, a subset of $\mathcal{Y}$, then each $y_0 \in \mathcal{D}$ for which $\delta J(y_0; v) = 0$, $\forall y_0 + v \in \mathcal{D}$ minimizes J on $\mathcal{D}$ [uniquely].*

III. STATEMENT OF THE PROBLEM AND ITS PROOF

It is well known that for an arbitrary functional J, the solution to the corresponding ELE is only a *necessary* condition for a local minimum. However, for functional that is [strictly] convex, Proposition 1 gives a surprisingly strong statement that not only the solution to ELE is a minimum but also a [unique] global minimum. In this paper, we prove the following proposition.

Proposition 2 *The GVF functional defined in (1) is [strictly] convex [when g and h are not both zero at any $\boldsymbol{x} \in \Omega$].*

PROOF. We need to show $J(\boldsymbol{v} + \boldsymbol{u}) - J(\boldsymbol{v}) \geq \delta J(\boldsymbol{v}; \boldsymbol{u})$ holds for $\forall \boldsymbol{u}$ such that $\boldsymbol{v} + \boldsymbol{u} \in \mathcal{D}$. By substituting the definition of GVF, we have

$$\begin{aligned}
&J(\boldsymbol{v} + \boldsymbol{u}) - J(\boldsymbol{v}) \\
&= \int_\Omega g|\nabla(\boldsymbol{v} + \boldsymbol{u})|^2 + h|\boldsymbol{v} + \boldsymbol{u} - \nabla f|^2 d\boldsymbol{x} \\
&\quad - \int_\Omega g|\nabla \boldsymbol{v}|^2 + h|\boldsymbol{v} - \nabla f|^2 d\boldsymbol{x} \\
&= \int_\Omega g(|\nabla \boldsymbol{v}|^2 + 2\nabla \boldsymbol{v} \cdot \nabla \boldsymbol{v} + |\nabla \boldsymbol{u}|^2) + h[|\boldsymbol{v} - \nabla f|^2 \\
&\quad + 2(\boldsymbol{v} - \nabla f) \cdot \boldsymbol{u} + |\boldsymbol{u}|^2] - g|\nabla \boldsymbol{v}|^2 - h|\boldsymbol{v} - \nabla f|^2 d\boldsymbol{x} \\
&= \int_\Omega g|\nabla \boldsymbol{u}|^2 + 2g\nabla \boldsymbol{v} \cdot \nabla \boldsymbol{u} + h|\boldsymbol{u}|^2 + 2h(\boldsymbol{v} - \nabla f) \cdot \boldsymbol{u} d\boldsymbol{x}.
\end{aligned}$$

Similarly, we can show

$$\begin{aligned}
J(\boldsymbol{v} + \varepsilon \boldsymbol{u}) - J(\boldsymbol{v}) &= \int_\Omega 2\varepsilon g \nabla \boldsymbol{v} \cdot \nabla \boldsymbol{u} + \varepsilon^2 g |\nabla u|^2 \\
&\quad + 2\varepsilon h(\boldsymbol{v} - \nabla f) \cdot \boldsymbol{u} + \varepsilon^2 h |\boldsymbol{u}|^2 d\boldsymbol{x}.
\end{aligned}$$

Therefore, the Gâteaux variation of GVF functional is given by

$$\delta J(\boldsymbol{v}; \boldsymbol{u}) = \lim_{\varepsilon \to 0} \frac{J(\boldsymbol{v} + \epsilon \boldsymbol{u}) - J(\boldsymbol{v})}{\varepsilon}$$

$$
\begin{aligned}
&= \lim_{\varepsilon\to 0}\int_\Omega 2g\nabla \boldsymbol{v}\cdot\nabla\boldsymbol{u} + \varepsilon g|\nabla\boldsymbol{u}|^2 \\
&\quad +2h(\boldsymbol{v}-\nabla f)\cdot\boldsymbol{u} + \varepsilon h|\boldsymbol{u}|^2 d\boldsymbol{x} \\
&= \int_\Omega 2g\nabla\boldsymbol{v}\cdot\nabla\boldsymbol{u} + 2h(\boldsymbol{v}-\nabla f)\cdot\boldsymbol{u}\, d\boldsymbol{x}.
\end{aligned}
$$

Thus, we have

$$
\begin{aligned}
&J(\boldsymbol{v}+\boldsymbol{u}) - J(\boldsymbol{v}) - \delta J(\boldsymbol{v};\boldsymbol{u}) \\
&= \int_\Omega g|\nabla\boldsymbol{u}|^2 + 2g\nabla\boldsymbol{v}\cdot\nabla\boldsymbol{u} + h|\boldsymbol{u}|^2 + 2h(\boldsymbol{v}-\nabla f)\cdot\boldsymbol{u} \\
&\quad -2g\nabla\boldsymbol{v}\cdot\nabla\boldsymbol{u} - 2h(\boldsymbol{v}-\nabla f)\cdot\boldsymbol{u}\, d\boldsymbol{x} \\
&= \int_\Omega g|\nabla\boldsymbol{u}|^2 + h|\boldsymbol{u}|^2 d\boldsymbol{x}.
\end{aligned}
$$

Since both g and h are nonnegative functions, we have

$$J(\boldsymbol{v}+\boldsymbol{u}) - J(\boldsymbol{v}) - \delta J(\boldsymbol{v};\boldsymbol{u}) \geq 0.$$

For g and h that do not reach zero at the same location, then the equality holds if and only if $\boldsymbol{u} = \boldsymbol{0}$. Hence, $J(\boldsymbol{v})$ is [strictly] convex [when g and h are not both zero at any $\boldsymbol{x}\in\Omega$]. □

The significance of convexity is seen in the following:

Proposition 3 *Let Ω be a domain in $\boldsymbol{R}^n$ and set $\mathcal{D} = W_2^2(\Omega)$, then each $\boldsymbol{v}\in\mathcal{D}$ satisfying GVF's ELE (see [6] for details in deriving the ELE)*

$$\nabla\cdot(g\nabla\boldsymbol{v}) = h(\boldsymbol{v}-\nabla f) \tag{2}$$

$$(\nabla\boldsymbol{v})\boldsymbol{N} = \boldsymbol{0} \text{ on } \partial\Omega, \tag{3}$$

where $\boldsymbol{N} = [N^1,\cdots,N^n]^T$ is the normal to the boundary $\partial\Omega$ and $(\nabla\boldsymbol{v})\boldsymbol{N} = [\sum_{i=1}^n \partial v^1/\partial x^i N^i,\cdots,\sum_{i=1}^n \partial v^n/\partial x^i N^i]^T$, minimizes GVF functional J on $\mathcal{D}$ [uniquely].

PROOF. As shown previously

$$\delta J(\boldsymbol{v};\boldsymbol{u}) = \int_\Omega 2g\nabla\boldsymbol{v}\cdot\nabla\boldsymbol{u} + 2h(\boldsymbol{v}-\nabla f)\cdot\boldsymbol{u}\, d\boldsymbol{x}. \tag{4}$$

We can integrate the first term by parts and use Green's theorem as follows:

$$
\begin{aligned}
\int_\Omega 2g\nabla\boldsymbol{v}\cdot\nabla\boldsymbol{u} &= \int_\Omega \nabla\cdot[2g(\nabla\boldsymbol{v})\boldsymbol{u}]d\boldsymbol{x} - \int_\Omega[\nabla\cdot(2g\nabla\boldsymbol{v})]\cdot\boldsymbol{u}d\boldsymbol{x} \\
&= \int_{\partial\Omega}[(2g\nabla\boldsymbol{v})\boldsymbol{u}]\cdot\boldsymbol{N}d\sigma - \int_\Omega[\nabla\cdot(2g\nabla\boldsymbol{v})]\cdot\boldsymbol{u}d\boldsymbol{x},
\end{aligned}
$$

where $d\sigma$ denotes the element of integration on $\partial\Omega$. Since

$$
\begin{aligned}
\int_{\partial\Omega}[(2g\nabla\boldsymbol{v})\boldsymbol{u}]\cdot\boldsymbol{N}d\sigma &= 2\int_{\partial\Omega} g[(\nabla\boldsymbol{v})^T\boldsymbol{N}]\cdot\boldsymbol{u}d\sigma \\
&= 2\int_{\partial\Omega} g[(\nabla\boldsymbol{v})\boldsymbol{N}]\cdot\boldsymbol{u} = 0
\end{aligned}
$$

(here $(\nabla\boldsymbol{v})^T = \nabla\boldsymbol{v}$ and $\boldsymbol{v}$ satisfies (3)), we have

$$
\begin{aligned}
\delta J(\boldsymbol{v};\boldsymbol{u}) &= -\int_\Omega[\nabla\cdot(2g\nabla\boldsymbol{v})]\cdot\boldsymbol{u}d\boldsymbol{x} + \int_\Omega 2h(\boldsymbol{v}-\nabla f)\cdot\boldsymbol{u}\, d\boldsymbol{x} \\
&= 2\int_\Omega[h(\boldsymbol{v}-\nabla f) - \nabla\cdot(g\nabla\boldsymbol{v})]\cdot\boldsymbol{u}d\boldsymbol{x}.
\end{aligned}
$$

Because $\boldsymbol{v}$ satisfies (2), $\delta J(\boldsymbol{v};\boldsymbol{u}) = 0$. Thus by Proposition 1 and Proposition 2, $\boldsymbol{v}$ minimizes J on $\mathcal{D}$ [uniquely]. □

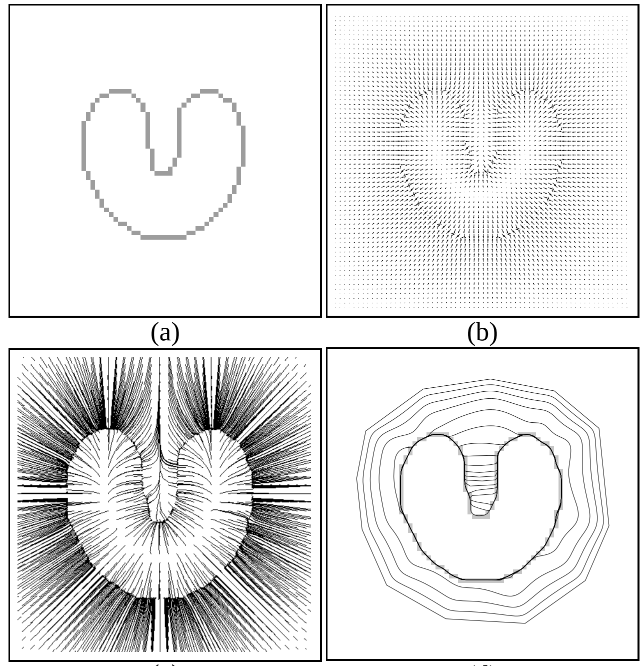

Fig. 1: (a) A line-drawing U-shape object. (b) The computed GVF field. (c) The streamline plot of the GVF field. (d) A 2-D deformable model with GVF as external forces progressively deforms from a polygon to the shape of the object.

IV. REMARKS

The weighting functions g and h, in practice, are always chosen in such a way that they are not both zero at the same location, hence following both Proposition 1 and Proposition 3, the resulting ELE always minimize the GVF functional uniquely on $\mathcal{D}$. Finally, we note that the existence of the solution to the GVF ELE has been proved in [7].

ACKNOWLEDGMENTS

The authors are particularly indebted to Wilson J. Rugh, who introduced the convexity analysis to the authors.

REFERENCES

[1] C. Xu and J. L. Prince, "Snakes, shapes, and gradient vector flow," *IEEE Trans. Imag. Proc.*, 7(3):359–369, 1998.

[2] C. Xu and J. L. Prince, "Generalized gradient vector flow external forces for active contours," *Signal Processing — An International Journal*, 71(2):131–139, 1998.

[3] T. McInerney and D. Terzopoulos, "Deformable models in medical image analysis: a survey," *Med. Imag. Anal.*, 1(2):91–108, 1996.

[4] J. L. Troutman, *Variational Calculus and Optimal Control*, Springer-Verlag, 1996.

[5] K. E. Gustafson, *Partial Differential Equations*, John Wiley & Sons, New York, 2 edition, 1987.

[6] C. Xu, *Deformable Models with Application to Human Cerebral Cortex Reconstruction from Magnetic Resonance Images*, PhD thesis, Department of Electrical and Computer Engineering, the Johns Hopkins University, January 1999.

[7] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, vol. 2, New York: Interscience, 1989.