[Supplementary] Distributed Cosegmentation via Submodular Optimization on Anisotropic Diffusion

Gunhee Kim[1] Eric P. Xing[1] Li Fei-Fei[2] Takeo Kanade[1]

[1] School of Computer Science, Carnegie Mellon University

[2] Computer Science Department, Stanford University

{gunhee,epxing,tk}@cs.cmu.edu feifeili@cs.stanford.edu

1. Proof of Theorem 1

Theorem 1 (Submodularity on Anisotropic Diffusion). Suppose that the system undergoes *linear anisotropic diffusion*. Let $u(x, t; \mathcal{S})$ be the temperature at position x at time t when identical heat sources are attached to $\mathcal{S}(\subset \Omega)$. Then, the following statements hold for $\forall x \in \Omega, \forall t \in [0, \infty]$.

(T1) $u(x, t; \emptyset) = 0$

(T2) $u(x, t; \mathcal{S})$ is *nondecreasing* and *submodular*.

Proof. Here we consider the discrete case where time and space are discretized; it is not difficult to draw the same conclusion for the continuous case. Without loss of generality, we assume that the source temperature is one and the environment temperature is zero. Then, the temperature can be interpreted as a *probability*. During the proof we drop t in the notation because the following arguments always hold for any t.

Note that the system is under *linear anisotropic* diffusion, which means that the system Ω and the diffusivity $D(x)$ including the dissipation diffusivity $z(x)$ are invariant for any t.

(T1) $u(x; \emptyset) = 0$ is obvious because without a source the system has zero temperature (*i.e.* the same temperature with that of environment).

(T2) $u(x; \mathcal{A})$ is *nondecreasing* (*i.e.* $u(x; \mathcal{A}) \leq u(x; \mathcal{B})$ for all $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{V}$) because the temperature of the system is always higher with more heat sources. Physically, it means the *energy conservation law*.

The $u(x; \mathcal{A})$ is *submodular* if Eq.(1) holds for all placements $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{V}$ and a new source $s \in \mathcal{V} \backslash \mathcal{B}$:

$$u(x; \mathcal{A} \cup \{s\}) - u(x; \mathcal{A}) \geq u(x; \mathcal{B} \cup \{s\}) - u(x; \mathcal{B}) \quad (1)$$

We shall prove the submodularity of u by induction on the distance $d(x, s)$. The induction proof consists of two steps, which are (a) **base step** showing that Eq.(1) holds for $d(x, s)$=0, and (b) **induction step** showing that if Eq.(1) holds for $d(x, s) \leq r$, then it is true for $d(x, s) \leq r+\delta r$ with a small $\delta r > 0$ as well.

(a) **Base step**: For x with $d(x, s) = 0$ (*i.e.* $x = s$), $u(x; \mathcal{A} \cup \{s\}) - u(x; \mathcal{A}) \geq u(x; \mathcal{B} \cup \{s\}) - u(x; \mathcal{B})$ because (i) $u(s; \mathcal{A} \cup \{s\}) = u(s; \mathcal{B} \cup \{s\}) = 1$ and (ii) $u(s; \mathcal{A}) \leq u(s; \mathcal{B})$ since $u(x; \mathcal{A})$ is nondecreasing for all $x \in \mathcal{V}$.

(b) **Induction step**: Suppose that for all x with $d(x, s) \leq r$, Eq.(1) holds. We need to show that Eq.(1) is true for all x' with $d(x', s) = r + \delta r$ with a small $\delta r > 0$ as well.

If the system undergoes diffusion, as shown in Eq.(2), the temperature at point x is represented by the weighted sum of the temperatures of its neighbors $\mathcal{N}(x)$ [1, 2]. It is based on the physical fact that the heat diffusion is driven by thermal non-equilibrium and converges to local energy balance.

$$u(x) = \sum_{p \in \mathcal{N}(x)} g(p) u(p) \qquad \text{for } \forall x \in \Omega \quad (2)$$

where $p \in \mathcal{N}(x)$ is a point of the neighbor set of x and $g(p)$ is a Kernel function describing how much the temperature at p ($u(p)$) contributes to the temperature at x ($u(x)$). $g(p)$ is the function of the diffusivity and the distance between p and x[1]. Therefore, $g(p)$ is invariant for any t under the *linear anisotropy* assumption (*i.e.* the system and the diffusivity are fixed for any t).

For a position x' with $d(x', s) = r + \delta r$, $\mathcal{N}(x')$ can be divided into two sets $\mathcal{P} = \{p | p \in \mathcal{N}(x'), d(p, s) \leq r\}$ and $\mathcal{Q} = \{q | q \in \mathcal{N}(x'), d(q, s) > r\}$. Therefore, $u(x'; \mathcal{A} \cup \{s\}) - u(x'; \mathcal{A}) \geq u(x'; \mathcal{B} \cup \{s\}) - u(x'; \mathcal{B})$ holds by Eq.(2) and induction hypotheses of (i) $u(p; \mathcal{A} \cup \{s\}) - u(p; \mathcal{A}) \geq u(p; \mathcal{B} \cup \{s\}) - u(p; \mathcal{B})$ for all $p \in \mathcal{P}$ and (ii) $u(q; \mathcal{A} \cup \{s\}) = u(q; \mathcal{A})$ and $u(q; \mathcal{B} \cup \{s\}) = u(q; \mathcal{B})$ for all $q \in \mathcal{Q}$.

[1] The simplest discrete form of Eq.(2) with a 2D regular grid is $u(i, j) = \big(u(i-1, j) + u(i+1, j) + u(i, j-1) + u(i, j+1)\big)/4$ with $x = (i, j)$. In this case, $\mathcal{N}(x) = \{(i, j-1), (i, j+1), (i-1, j), (i+1, j)\}$ and $g(p) = 1/4$ for $\forall p \in \mathcal{N}(x)$. In a more accurate discretization [1], the Gaussian Kernel is used: $g(p) = \exp\big(-(x-p)^T D(p)(x-p)/\sigma_p\big)$ where σ_p is a normalization constant so that $\sum_{p \in \mathcal{N}(x)} g(p) = 1$.

References

[1] D. Tschumperle and R. Deriche. Vector-Valued Image Regularization with PDE's : A Common Framework for Different Applications. *IEEE PAMI*, 27:506–517, 2005. 1

[2] J. Weickert. *Anisotropic Diffusion in Image Processing*. ECMI Series, Teubner-Verlag, 1998. 1