Supplementary Material: Coarse-to-Fine Segmentation With Shape-Tailored Continuum Scale Spaces

Naeemullah Khan[1], Byung-Woo Hong[2], Anthony Yezzi[3], and Ganesh Sundaramoorthi[1]
[1]KAUST, Saudi Arabia [2] Chung-Ang University, Korea [3] Georgia Tech, USA
{naeemullah.khan,ganesh.sundaramoorthi}@kaust.edu.sa, hong@cau.ac.kr, ayezzi@ece.gatech.edu

1. Proofs of Lemmas and Propositions

1.1. Proofs Regarding Constrained Optimization Problem

Lemma 1 *Suppose $I : \mathbb{R}^2 \to \mathbb{R}$ and $a = avg(I) = avg(u(t, \cdot))$. Then*

$$E = \int_0^\infty \int_{\mathbb{R}^2} |u(t,x) - a|^2 \,\mathrm{d}x \,\mathrm{d}t = \int_{\mathbb{R}^2} |H(\omega)\hat{I}(\omega)|^2 \,\mathrm{d}\omega, \text{ where } H(\omega) = \frac{1}{\sqrt{2}|\omega|}, \tag{1}$$

where $\hat{I}$ denotes the Fourier transform, and ω denotes frequency.

Proof 1 *Taking the Fourier transform of the Heat Equation:*

$$\begin{cases} \partial_t u(t,x) = \Delta u(t,x) & x \in \mathbb{R}^2 \\ u(0,x) = I(x) & t = 0 \end{cases}$$

yields:

$$\partial_t \hat{u}(t,\omega) = (i\omega)\cdot(i\omega)\hat{u}(t,\omega) = -|\omega|^2 \hat{u}(t,\omega),$$

where $\hat{u}(t,\omega)$ is the Fourier transform of u. Solving this differential equation yields

$$\hat{u}(t,\omega) = e^{-|\omega|^2 t}\hat{I}(\omega).$$

We note that $a = 0$ when $I \in \mathbb{L}^2$ since $\hat{I}(0) = \int_{\mathbb{R}^2} I(x)\,\mathrm{d}x$ is finite. Then by Parseval's Theorem,

$$E = \int_0^\infty \int_{\mathbb{R}^2} |u(t,x) - a|^2 \,\mathrm{d}x\,\mathrm{d}t = \int_0^\infty \int_{\mathbb{R}^2} |\hat{u}(t,\omega)|^2 \,\mathrm{d}\omega\,\mathrm{d}t \tag{2}$$

$$= \int_{\mathbb{R}^2} \int_0^\infty e^{-2|\omega|^2 t}\,\mathrm{d}t \cdot |\hat{I}(\omega)|^2 \,\mathrm{d}\omega \tag{3}$$

$$= \int_{\mathbb{R}^2} -\frac{1}{2|\omega|^2} e^{-2|\omega|^2 t}\Big|_{t=0}^{t=+\infty} |\hat{I}(\omega)|^2 \,\mathrm{d}\omega = \int_{\mathbb{R}^2} \frac{1}{2|\omega|^2}|\hat{I}(\omega)|^2 \,\mathrm{d}\omega \tag{4}$$

$$= \int_{\mathbb{R}^2} |A(\omega)\hat{I}(\omega)|^2 \,\mathrm{d}\omega \tag{5}$$

where $A(\omega) = 1/(\sqrt{2}|\omega|)$.

Lemma 2 (PDE for Lagrange Multiplier λ) *The Lagrange multiplier λ satisfies the following Heat Equation with forcing term, evolving backwards in time:*

$$\begin{cases} \partial_t \lambda(t,x) + \Delta\lambda(t,x) = f'(u(t,x)) & x \in R \times [0,T] \\ \nabla\lambda(t,x)\cdot N = 0 & x \in \partial R \times [0,T] \\ \lambda(T,x) = 0 & x \in R \end{cases}. \tag{6}$$

Proof 2 *We define*

$$E(R,u,\lambda) = \int_R \int_0^T f(u)\,\mathrm{d}x\,\mathrm{d}t + \int_R \int_0^T \left(\nabla\lambda \cdot \nabla u + \lambda \cdot u_t\right)\,\mathrm{d}x\,\mathrm{d}t. \tag{7}$$

Integrating by parts, we have that

$$E = \int_{R\times[0,T]} \left[f(u) - (\partial_t \lambda + \Delta\lambda)u\right]\,\mathrm{d}x\,\mathrm{d}t \tag{8}$$

$$+ \int_R \lambda u|_{t=0}^{t=T}\,\mathrm{d}x + \int_{\partial R\times[0,T]} (\nabla\lambda \cdot N)u\,\mathrm{d}s(x)\,\mathrm{d}t, \tag{9}$$

where $\mathrm{d}s$ *denotes the arc-length measure of* ∂R*, and* N *is the unit outward normal of* ∂R*. Differentiating* E *in the direction (perturbation)* $\tilde{u}$ *of* u *evaluated at* u *yields*

$$\mathrm{d}E(u)\cdot\tilde{u} = \int_{R\times[0,T]} \left[f'(u) - (\partial_t \lambda + \Delta\lambda)\right]\tilde{u}\,\mathrm{d}x\,\mathrm{d}t \tag{10}$$

$$+ \int_R \lambda \tilde{u}|_{t=0}^{t=T}\,\mathrm{d}x + \int_{\partial R\times[0,T]} (\nabla\lambda \cdot N)\tilde{u}\,\mathrm{d}s(x)\,\mathrm{d}t. \tag{11}$$

Note that $\tilde{u}(0) = 0$ *since* $u(0) = I$ *is fixed and thus may not be perturbed. We may choose* $\nabla\lambda \cdot N = 0$ *on* ∂R *and* $\lambda(T) = 0$*. We are interested in* u *such that* $\mathrm{d}E(u)\cdot\tilde{u} = 0$ *for all* $\tilde{u}$*. This yields the condition that*

$$\begin{cases} \partial_t \lambda(t,x) + \Delta\lambda(t,x) = f'(u(t,x)) & x \in R\times[0,T] \\ \nabla\lambda(t,x)\cdot N = 0 & x \in \partial R \times [0,T] \\ \lambda(T,x) = 0 & x \in R \end{cases}. \tag{12}$$

Lemma 3 (Lagrange Multiplier λ) *The solution of* (6) *can be written as*

$$\lambda(t,x) = -\int_t^T F(s-t,x;s)\,\mathrm{d}s. \tag{13}$$

where $F(\cdot,\cdot;s) : [0,T]\times R \to \mathbb{R}$ *is the solution of the forward heat equation* (**??**) *with zero forcing and initial condition* $f'(u)$ *evaluated at time* s*, i.e.,*

$$\begin{cases} \partial_t F(t,x;s) - \Delta F(t,x;s) = 0 & x \in R\times[0,T] \\ \nabla F(t,x;s)\cdot N = 0 & x \in \partial R \times [0,T] \\ F(0,x;s) = f_u(s,u(s,x)) & x \in R \end{cases}. \tag{14}$$

In the case that $f(t,u) = (u-a)^2 w(t)$*,* λ *can be expressed as*

$$\lambda(t,x) = -2\int_t^T (u(2s-t,x) - a)w(s)\,\mathrm{d}s. \tag{15}$$

Proof 3 *To express the solution to the above equation in a more convenient form, we may use Duhamel's Principle. The latter states that a linear PDE with forcing term* $\delta(t-s)$ *is equivalent to the same PDE with zero forcing and initial condition at* s *of* 1*. We may express the forcing term as* $\int_{[0,T]} f(u(x,s))\delta(s-t)\,\mathrm{d}s$*, and thus combining linearity of the PDE with Duhamel's Principle yields that*

$$\lambda(t,x) = -\int_t^T F(s-t,x;s)\,\mathrm{d}s, \tag{16}$$

i.e., it is the sum of solutions of the PDE with zero forcing and initial condition $f(u(x,s))$ *at time* s*, specifically,*

$$\begin{cases} \partial_t F(t,x;s) - \Delta F(t,x;s) = 0 & (t,x) \in [0,T]\times R \\ \nabla F(t,x;s)\cdot N = 0 & (t,x) \in [0,T]\times \partial R \\ F(0,x;s) = f(u(s,x)) & x \in R \end{cases}. \tag{17}$$

In the case that $f(u) = (u-a)^2$ then $f'(u) = 2(u-a)$, the PDE for F becomes

$$\begin{cases} \partial_t F(t,x;s) = \Delta F(t,x;s) & x \in R \times [0,T] \\ \nabla F(t,x;s) \cdot N = 0 & x \in \partial R \times [0,T] \\ F(0,x;s) = 2(u(s,x) - a) & x \in R \end{cases}, \tag{18}$$

which is the forward Heat Equation with initial condition being the solution of the same Heat Equation evaluated at time s. By the semi-group property of the Heat Equation, we have that

$$F(t,x;s) = 2(u(s+t,x) - a), \tag{19}$$

and therefore using (16),

$$\lambda(t,x) = -2\int_t^T (u(s+s-t,x) - a)\,\mathrm{d}s = -2\int_t^T (u(2s-t,x) - a)\,\mathrm{d}s. \tag{20}$$

Proposition 1 *The gradient of E with respect to the boundary ∂R can be expressed as*

$$\nabla_{\partial R} E = \int_0^T \left[f(u) + \nabla\lambda \cdot \nabla u + \lambda \partial_t u\right] \mathrm{d}t \cdot N, \tag{21}$$

where N is the normal vector to ∂R.

Proof 4 *To compute the gradient of E, we compute the gradient of E in (7) with respect to ∂R treating λ and u independent of R as in the theory of Lagrange multipliers. In this case, this is just a classical result in the calculus of variations (e.g., [1]), in particular the integrand (with respect to R) is multiplied by the outward normal along ∂R to obtain the gradient:*

$$\nabla_{\partial R} E = \int_0^T \left[f(u) + \nabla\lambda \cdot \nabla u + \lambda \partial_t u\right] \mathrm{d}t \cdot N. \tag{22}$$

1.2. Weighting Functions

We now specialize the results in the previous section and consider

$$E(R) = \int_R \int_0^T (u(t,x) - a)^2 w(t)\,\mathrm{d}t\,\mathrm{d}x,$$

where u satisfies the heat equation. The gradient is

$$\nabla_{\partial R} E = \int_0^T \left[(u-a)^2 w + \nabla\lambda \cdot \nabla u + \lambda \partial_t u\right] \mathrm{d}t$$

where

$$\begin{cases} \partial_t \lambda(t,x) + \Delta\lambda(t,x) = 2(u(t,x) - a)w(t) & x \in R \times [0,T] \\ \nabla\lambda(t,x) \cdot N = 0 & x \in \partial R \times [0,T] \\ \lambda(T,x) = 0 & x \in R \end{cases}$$

The Lagrange multiplier satisfies

$$\lambda(t,x) = -\int_t^T F(s-t,x;s)\,\mathrm{d}s$$

where

$$\begin{cases} \partial_t F(t,x;s) = \Delta F(t,x;s) & x \in R \times [0,T] \\ \partial F(t,x;s) \cdot N = 0 & x \in \partial R \times [0,T] \\ F(0,x;s) = 2(u(s,x) - a)w(s) & x \in R \end{cases},$$

which has solution

$$F(t,x;s) = 2(u(s+t,x) - a)w(s).$$

Therefore,

$$\begin{aligned}
\lambda(t,x) &= -2\int_t^T (u(2s-t,x)-a)w(s)\,\mathrm{d}s \\
&= -\int_t^{2T-t} (u(\tau,x)-a)w((\tau+t)/2)\,\mathrm{d}\tau.
\end{aligned}$$

Note

$$\begin{aligned}
\partial_t\lambda(t,x) &= 2(u(t,x)-a)w(t) + 2\int_t^T \Delta u(2s-t,x)w(s)\,\mathrm{d}s \\
&= 2(u(t,x)-a)w(t) + \int_t^T \partial_s u(2s-t,x)w(s)\,\mathrm{d}s \\
&= 2(u(t,x)-a)w(t) + u(2T-t,x)w(T) - u(t,x)w(t) - \int_t^T u(2s-t,x)w'(s)\,\mathrm{d}s \\
&= (u(t,x)-2a)w(t) + u(2T-t,x)w(T) - \frac{1}{2}\int_t^{2T-t} u(\tau,x)w'((\tau+t)/2)\,\mathrm{d}\tau
\end{aligned}$$

The gradient becomes

$$\begin{aligned}
\nabla_{\partial R}E &= \int_0^T [(u_t-a)^2 w_t - u_t(u_t-2a)w_t - u_{2T-t}w_T]\,\mathrm{d}t \\
&\quad + \frac{1}{2}\int_0^T\int_t^{2T-t} u_t u_\tau w'((\tau+t)/2)\,\mathrm{d}\tau\,\mathrm{d}t + \int_0^T \nabla\lambda\cdot\nabla u\,\mathrm{d}t + \lambda_0 u_0 \\
&= \int_0^T a^2 w_t - u_{2T-t}w_T\,\mathrm{d}t + \lambda_0 u_0 \\
&\quad + \frac{1}{2}\int_0^T\int_t^{2T-t} u_t u_\tau w'((\tau+t)/2)\,\mathrm{d}\tau\,\mathrm{d}t - \int_0^T\int_t^{2T-t} \nabla u_\tau\cdot\nabla u_t w((\tau+t)/2)\,\mathrm{d}\tau\,\mathrm{d}t \\
&= a^2\int_0^T w_t\,\mathrm{d}t - w_T\int_T^{2T} u_t\,\mathrm{d}t + \lambda_0 u_0 \\
&\quad + \frac{1}{4}\int_0^{2T}\int_0^{2T-t} u_t u_\tau w'((\tau+t)/2)\,\mathrm{d}\tau\,\mathrm{d}t - \frac{1}{2}\int_0^{2T}\int_0^{2T-t} \nabla u_\tau\cdot\nabla u_t w((\tau+t)/2)\,\mathrm{d}\tau\,\mathrm{d}t
\end{aligned}$$

1.2.1 Exponential weight

Let

$$w(t) = e^{\alpha t},$$

for any $\alpha\in\mathbb{R}$. Then

$$\begin{aligned}
w((\tau+t)/2) &= e^{\alpha t/2}e^{\alpha\tau/2} \\
w'((\tau+t)/2) &= \alpha e^{\alpha t/2}e^{\alpha\tau/2} \\
a^2\int_0^T w_t\,\mathrm{d}t &= \frac{a^2}{\alpha}(e^{\alpha T}-1)
\end{aligned}$$

$$\lambda_0 u_0 = -u_0\int_0^{2T}(u(\tau,x)-a)e^{\alpha\tau/2}\,\mathrm{d}\tau = -u_0\int_0^{2T} u_\tau e^{\alpha\tau/2}\,\mathrm{d}\tau + \frac{2a}{\alpha}(e^{\alpha T}-1)u_0.$$

The gradient is

$$\nabla_{\partial R}E = \frac{a}{\alpha}(e^{\alpha T}-1)(a+2u_0) - u_0\int_0^{2T} u_\tau e^{\alpha\tau/2}\,\mathrm{d}\tau - e^{\alpha T}\int_T^{2T} u_t\,\mathrm{d}t$$
$$+\frac{\alpha}{4}\int_0^{2T}\int_0^{2T-t} u_t e^{\alpha t/2}\cdot u_\tau e^{\alpha\tau/2}\,\mathrm{d}\tau\,\mathrm{d}t - \frac{1}{2}\int_0^{2T}\int_0^{2T-t}\nabla[u_t e^{\alpha t/2}]\cdot\nabla[u_\tau e^{\alpha\tau/2}]\,\mathrm{d}\tau\,\mathrm{d}t.$$

Assuming T is large and setting $U=\int_0^{2T} u_\tau e^{\alpha\tau/2}\,\mathrm{d}\tau$, we have

$$\nabla_{\partial R}E = \frac{a}{\alpha}(e^{\alpha T}-1)(a+2u_0) - aTe^{\alpha T} - u_0 U + \frac{\alpha}{4}U^2 - \frac{1}{2}|\nabla U|^2.$$

We can calculate U as

$$\Delta U = \int_0^{2T}\partial_\tau u_\tau e^{\alpha\tau/2}\,\mathrm{d}\tau = u_{2T}e^{\alpha T} - u_0 - (\alpha/2)U \approx ae^{\alpha T} - u_0 - (\alpha/2)U$$

or

$$\begin{cases} -U-\beta\Delta U = \beta(u_0 - ae^{\alpha T}) & x\in R \\ \nabla U\cdot N = 0 & x\in\partial R \end{cases}$$

where $\beta = \frac{2}{\alpha}$.

1.3. Exponential weight (negative exponent)

We set

$$w(t) = e^{-\alpha t},$$

for $\alpha > 0$, and take $T\to\infty$.

We can just substitute $-\alpha$ for α in the previous section's result, and then take the limit as $T\to\infty$. This gives

$$\nabla_{\partial R}E = \frac{a}{\alpha}(a+2u_0) - u_0 U + \frac{\alpha}{4}U^2 - \frac{1}{2}|\nabla U|^2,$$

where

$$\begin{cases} U - \frac{2}{\alpha}\Delta U = \frac{2}{\alpha}u_0 & x\in R \\ \nabla U\cdot N = 0 & x\in\partial R \end{cases}.$$

1.4. Truncated Uniform Weight

We consider

$$w(t) = \mathbf{1}_{[T_m,T]}(t),$$

and let $t_m = \max\{t, T_m\}$.

We see that

$$\begin{aligned}\partial_t\lambda(t,x) &= 2(u(t,x)-a)w(t) + \int_t^T \partial_s u(2s-t,x)w(s)\,\mathrm{d}s \\ &= 2(u(t,x)-a)w(t) + \int_{t_m}^T \partial_s u(2s-t,x)\,\mathrm{d}s \\ &= 2(u(t,x)-a)w(t) + u(2T-t,x) - u(t_m,x).\end{aligned}$$

So

$$(\partial_t\lambda_t)u_t = 2u_t(u_t-a)w_t + u_{2T-t}u_t - u_{t_m}u_t$$

Also,

$$\lambda_t = -\int_{t_m}^{2T-t}(u_\tau - a)\,\mathrm{d}\tau,$$

and

$$\lambda_0 = -\int_{T_m}^{2T} (u_\tau - a)\,\mathrm{d}\tau.$$

$$\begin{aligned}
\int_0^T [(u_t-a)^2 w_t - (\partial_t \lambda_t) u_t]\,\mathrm{d}t &= \int_0^T w_t[(u_t-a)^2 - 2u_t(u_t-a)] - u_{2T-t}u_t + u_{t_m}u_t\,\mathrm{d}t \\
&= \int_{T_m}^T (a^2 - u_t^2)\,\mathrm{d}t - \int_0^T u_{2T-t}u_t - u_{t_m}u_t\,\mathrm{d}t \\
&= \int_{T_m}^T (a^2 - u_t^2)\,\mathrm{d}t - \int_0^T u_{2T-t}u_t\,\mathrm{d}t + u_{T_m}\int_0^{T_m} u_t\,\mathrm{d}t + \int_{T_m}^T u_t^2\,\mathrm{d}t \\
&= a^2(T - T_m) - \int_0^T u_{2T-t}u_t\,\mathrm{d}t + u_{T_m}\int_0^{T_m} u_t\,\mathrm{d}t.
\end{aligned}$$

Also,

$$\int_0^T \nabla u_t \cdot \nabla \lambda_t\,\mathrm{d}t = -\int_0^T \int_{t_m}^{2T-t} \nabla u_t \cdot \nabla u_\tau\,\mathrm{d}\tau\,\mathrm{d}t$$

Let us assume T is large so that $u_t \approx a$ for $t \geq T$. Then

$$\lambda_0 = -\int_{T_m}^{T} (u_\tau - a)\,\mathrm{d}\tau$$

$$\int_0^T [(u_t-a)^2 w_t - (\partial_t \lambda_t) u_t]\,\mathrm{d}t = a^2(T - T_m) - a\int_0^T u_t\,\mathrm{d}t + u_{t_m}\int_0^{T_m} u_t\,\mathrm{d}t$$

$$\int_0^T \nabla u_t \cdot \nabla \lambda_t\,\mathrm{d}t = -\int_0^T \int_{t_m}^{T} \nabla u_t \cdot \nabla u_\tau\,\mathrm{d}\tau\,\mathrm{d}t = -\frac{1}{2}\left|\int_0^T \nabla u_t\,\mathrm{d}t\right|^2 + \frac{1}{2}\left|\int_0^{T_m} \nabla u_t\,\mathrm{d}t\right|^2.$$

Define

$$U_T = \int_0^T u_t\,\mathrm{d}t, \quad U_{T_m} = \int_0^{T_m} u_t\,\mathrm{d}t;$$

then

$$\lambda_0 = -(U_T - U_{T_m}) + a(T - T_m)$$

$$\int_0^T [(u_t-a)^2 w_t - (\partial_t \lambda_t) u_t]\,\mathrm{d}t = a^2(T - T_m) - aU_T + u_{T_m}U_{T_m}$$

$$\int_0^T \nabla u_t \cdot \nabla \lambda_t\,\mathrm{d}t = -\frac{1}{2}|\nabla U_T|^2 + \frac{1}{2}|\nabla U_{T_m}|^2.$$

Therefore

$$\nabla_{\partial R} E \approx a(T - T_m)(u_0 + a) - (u_0 + a)U_T + (u_{T_m} + a)U_{T_m} - \frac{1}{2}|\nabla U_T|^2 + \frac{1}{2}|\nabla U_{T_m}|^2.$$

Letting $T_m = 0$ gives Eqn 12 in the paper.

2. Discretization of the Poisson Equation

We show how to discretize Eqn (11) in the paper. The discretization of the Laplacian is

$$-\Delta U(x) = -\sum_{y \sim x} U(y) - U(x), \tag{23}$$

where $y \sim x$ indicates that y is a 4-neighbor of x. Discretizing the boundary condition $\nabla U(x) \cdot N = U(y) - U(x) = 0$, when $y \sim x$, $y \notin R$, and substituting it above, and then discreting Eqn 12 we have that

$$U(x) - T \cdot \sum_{y \sim x, y \in w_i(R_i)} U(y) - U(x) = T u_0(x). \tag{24}$$

This can now be directly used in any iterative solver (e.g., conjugate gradient or multigrid).

Real-World Texture Dataset

	Contour		Region metrics					
	F-meas.		GT-cov.		Rand. Index		Var. Info.	
	ODS	OIS	ODS	OIS	ODS	OIS	ODS	OIS
5×5	0.60	0.60	0.91	0.91	0.91	0.91	0.45	0.45
4×4	0.60	0.60	0.91	0.91	0.91	0.91	0.45	0.45
3×3	0.60	0.60	0.92	0.92	0.92	0.92	0.45	0.45

Table 1. **Sensitivity to initialization**: Varying the box tessellation initialization for the segmentation doesn't have a significant effect on the results of the segmentation

3. Additional Results

3.1. Visual Results on Texture Segmentation

We provide additional visual comparisons on samples from the Real World Texture Dataset in Figures 1 and 2. We compare our new approach using a continuum of scales (ExpNeg) with STLD (using discrete scales). See paper for more descriptions of the methods. As the figure shows our approach captures the segmentation more accurately than STLD.

3.2. Quantitative Analysis of Robustness to Initialization

We quantitatively show that our method is not sensitive to initialization. See Table 1, which are the results on the Real World Texture Dataset as we vary the number of boxes in the standard box tessellation from a 3×3 grid of boxes to 5×5. Note that the boxes are placed to cover the image. Results indicate very little difference in the final segmentations, which indicates that the parameters of the box tessellation doesn't affect the results much.

3.3. Visual Results Comparing Discrete and Continuum Scale Spaces

Can using a large number of discrete scales (in STLD) approximate the continuum approach? We show a visual result that indicates that the coarse-to-fine property, i.e., the overwhelming preference to segment the coarse structure before the fine scale details, is a property of the continuum scale space defined by the heat equation. We thus compare our approach to the discrete scale space approach (STLD) where each discrete scale is defined by an equation similar to the Poisson equation. We experiment with a fixed range of scales (0 to 40) for the continuum scale space, using the Uniform weighting. We fix the same range for the discrete scale space, but choose a varying number of scales within that range (5 to 500 scales). Results are shown in Figure 3. They show that even as the discrete scales are chosen large to approach the continuum, the approach seems to be sensitive to fine scale details. The continuum approach appears to be much less sensitive to fine details, and this seems to be a property of the heat equation. Note that in practice it would never be feasible to calculate beyond 5 scales in the discrete approach, as one would have to solve explicitly for each of those scales, a time-consuming process.

What about choosing a larger min scale in STLD? While one can increase the minimum scale considered from the native scale to a larger (more blurred) scale, that requires an extra parameter choice, which varies in general with the image. Some results are shown in Figure 4. Notice results in STLD improve as the minimum scale is increased, but the minimum scale could vary based on the scale content of the image. Our continuum approach does not require a choice for the minimum scale, providing added convenience.

3.4. Visual Results on FBMS

We provide additional results on the samples of FBMS59 dataset. We also enclose some videos. Some visual results are seen in Figure 5. We compare our continuum scale space formulation of motion segmentation to a single scale formulation at the native scale (non-SS).

As can be seen, small errors from a few frames accumulate and the results degrade as errors are propagated through the sequence in the non-SS approach. In contrast, our continuum scale space approach is less sensitive to fine scale details, and thus less errors are propagated.

References

[1] Zhu, S.C., Yuille, A.: Region competition: Unifying snakes, region growing, and bayes/mdl for multiband image segmentation. Pattern Analysis and Machine Intelligence, IEEE Transactions on **18**(9) (1996) 884–900 3

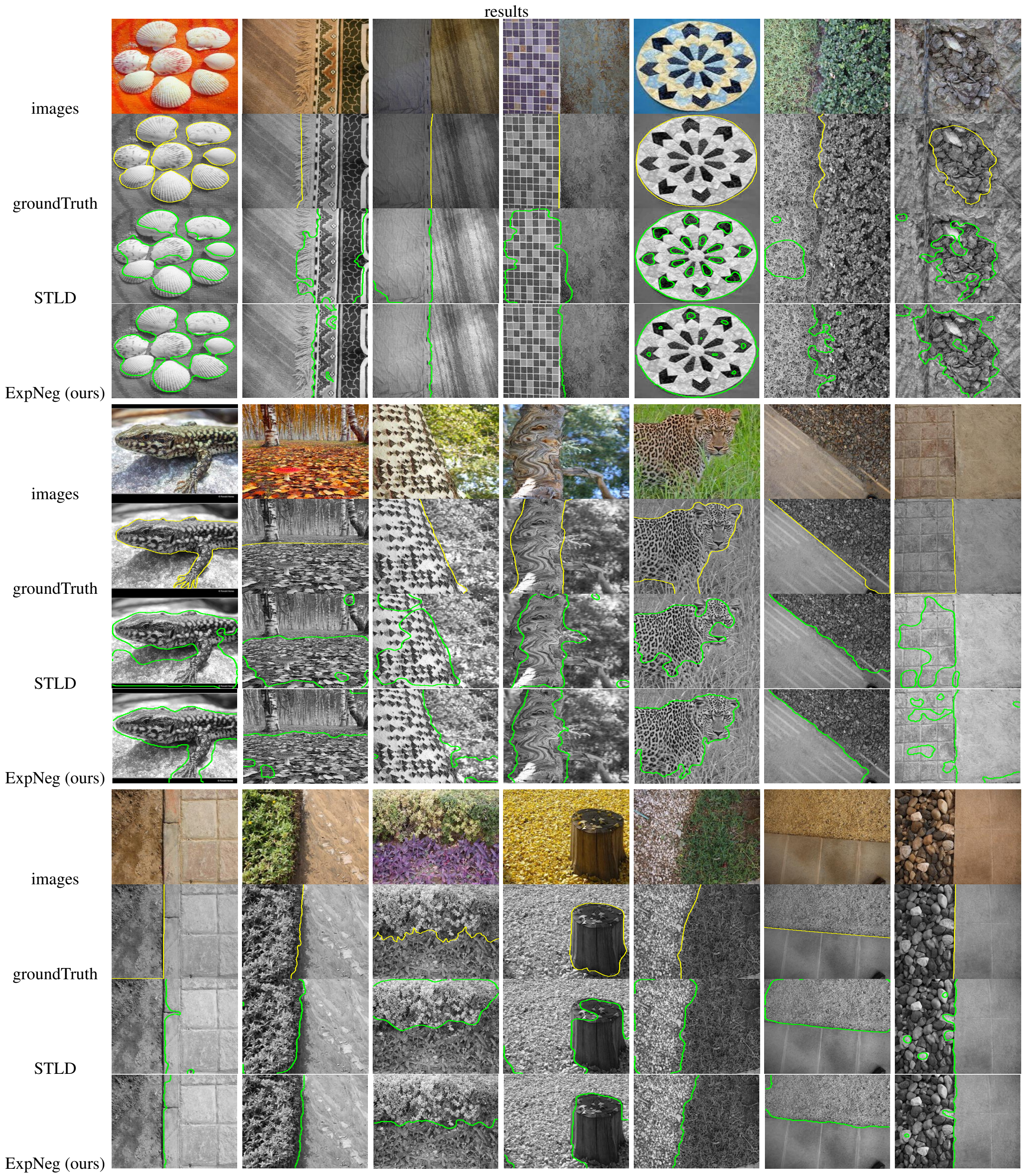

Figure 1. **Sample visual results on representative images from the Real World Texture Dataset**: The change of energy to integrate over a continuum of scales (ExpNeg) is generally less sensitive to clutter than using an energy that contains only a few scales (STLD).

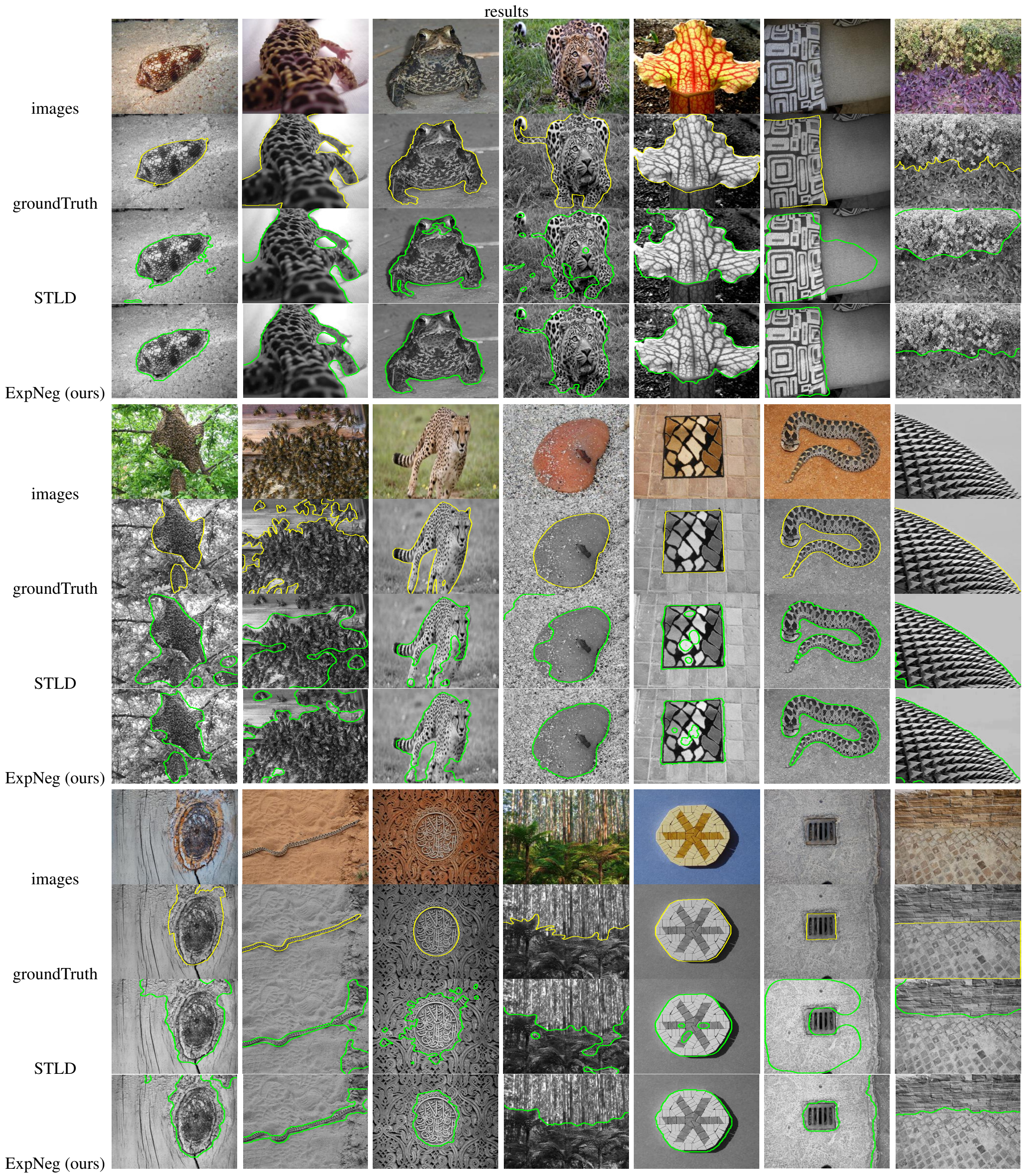

Figure 2. **Sample visual results on representative images from the Real World Texture Dataset**: The change of energy to integrate over a continuum of scales (ExpNeg) is generally less sensitive to clutter than using an energy that contains only a few scales (STLD).

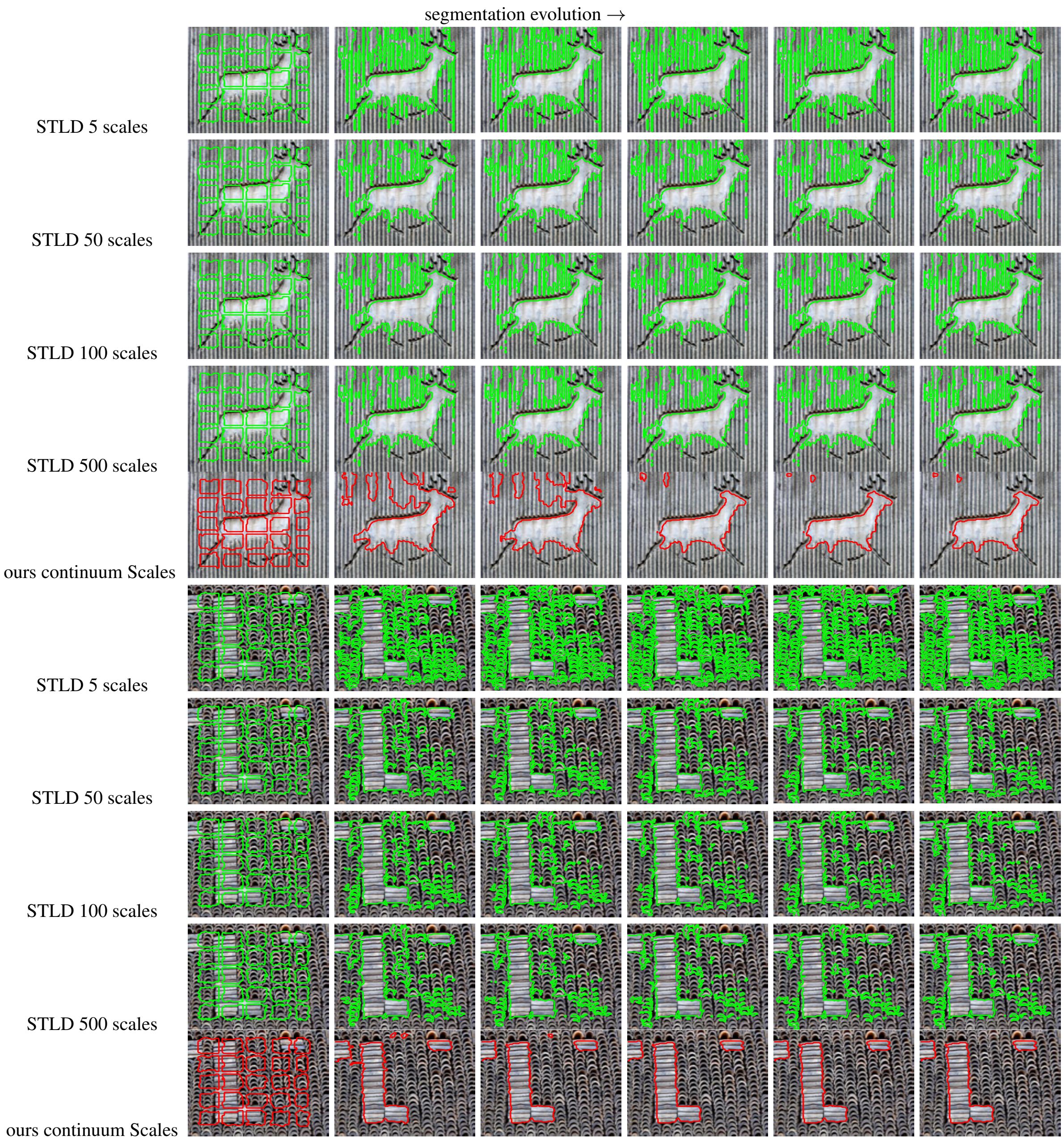

Figure 3. **Coarse to fine property of Continuum of Scales vs Discrete Scales (STLD).** The experiment shows that even if we increase the number of scales with STLD (even to an exorbitant number), STLD is still sensitive to the fine details. Indeed, it reacts immediately to the fine scale and does not exhibit the coarse-to-fine property seen with the continuum approach.

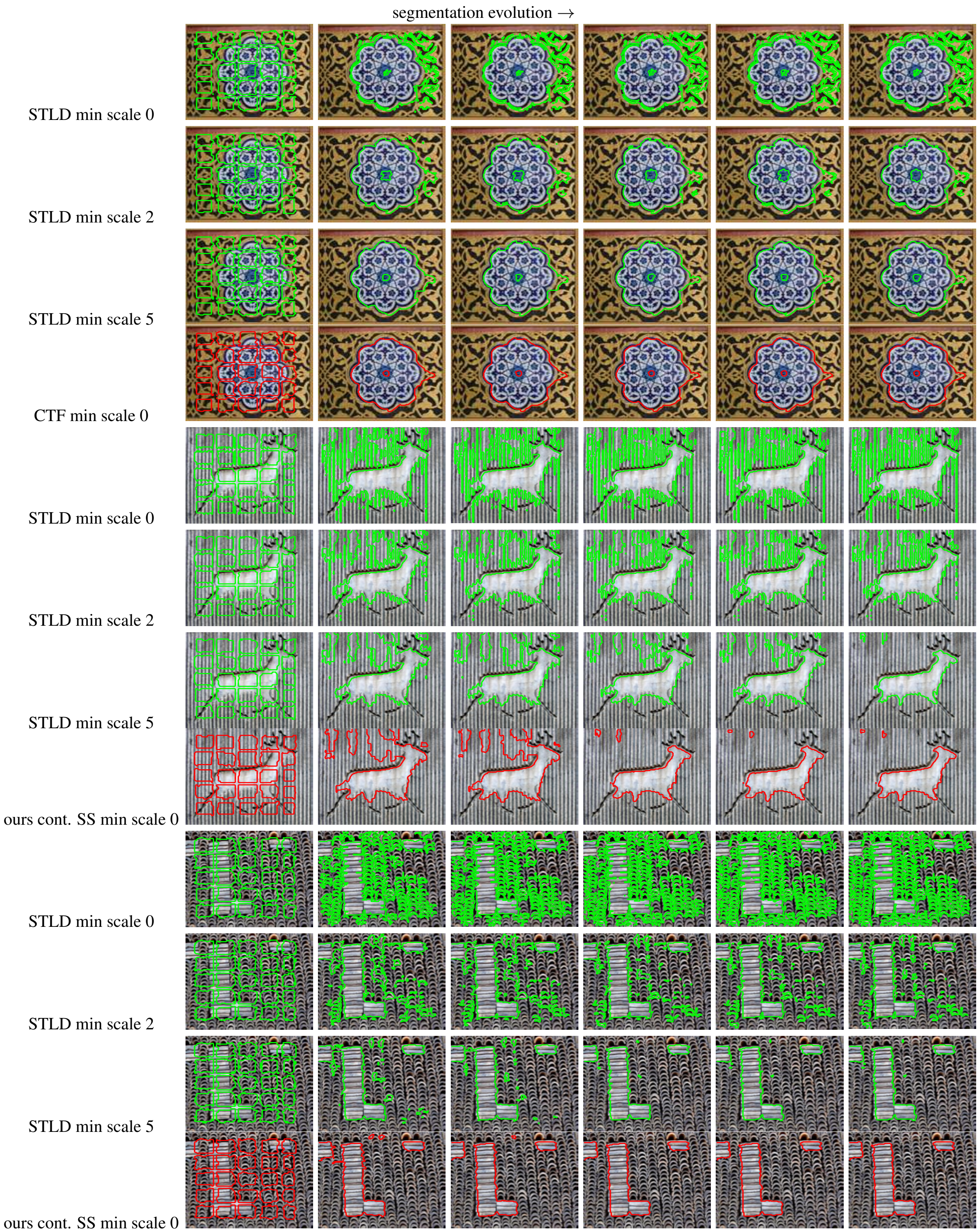

Figure 4. **Comparison of STLD with varying min scale and the continuum approach (ours).** Since STLD doesn't have the coarse to fine property, the discrete scales have to be adjusted so that they match the scales of the image to get a good segmentation. The continuum approach, since it possess a coarse to fine property, can achieve good results without tuning the minimum scale.

Figure 5. **Sample visual results on representative sequences for the FBMS-59 dataset** (segmented objects in purple and red). The continuum of scales (ours) is less sensitive to clutter than using an energy that contains only considers the native scale of the image (non-SS).