SUPPLEMENTARY MATERIALS: Dense Scale Selection Over Space, Time and Space-Time*

Tony Lindeberg†

SM1. Detailed analysis of scale calibration for dense spatial scale selection and when using spatial post-smoothing. In the treatment of dense spatial scale selection in Section 2 in the main article, we analysed the scale selection properties obtained from detecting local extrema over scale of the scale-normalized spatial quasi quadrature measure (8)

$$\mathcal{Q}_{(x,y),\Gamma-norm} L = \frac{s\,(L_x^2 + L_y^2) + C_s\, s^2\,(L_{xx}^2 + 2L_{xy}^2 + L_{yy}^2)}{s^{\Gamma_s}} \tag{SM1}$$

in Section 2.2. Specifically, for the case of using only phase compensation while no post-smoothing to reduce the phase dependency of the spatial scale estimates, the spatial scale estimates are for a 1-D sine wave with angular frequency ω_0 centered around the spatial scale level (30)

$$\hat{s} = \sqrt{\hat{s}_1\, \hat{s}_2} = \frac{\sqrt{(1-\Gamma_s)(2-\Gamma_s)}}{\omega_0^2} \tag{SM2}$$

By the notion of scale calibration described in Section 2.4, it was proposed that these scale estimates can be calibrated by multiplication with a uniform scale calibration factor to be either:

(i) equal to the scale estimate $\hat{s} = s_0$ obtained by applying the regular scale normalized Laplacian $\nabla^2_{norm} L = s\,(L_{xx} + L_{yy})$ or the scale-normalized determinant of the Hessian $\det \mathcal{H}_{norm} L = s^2\,(L_{xx} L_{yy} - L_{xy}^2)$ at the center of a Gaussian blob of any spatial extent s_0.

(ii) equal to the scale estimate $\hat{s} = \sqrt{2}/\omega_0^2$ corresponding to the geometric average of the scale estimates obtained for a sine wave of any angular frequency ω_0 when $\Gamma_s = 0$.

In this supplement section, we describe in more detail how such scale calibration can be performed when also using spatial post-smoothing.

SM1.1. Influence of the post-smoothing scale. When applying post-smoothing, the variability in the spatial scale estimates decreases with the relative integration scale parameter c (see the third, fourth and fifth rows in Figure 3). The minimum scale estimates are assumed at the spatial points at which $\mathcal{Q}_{(x,y),\Gamma-norm} L$ only responds to first-order information, whereas the maximum scale estimates are assumed at the points at which $\mathcal{Q}_{(x,y),\Gamma-norm} L$ only responds to second-order information.

*Supplementary material for SIIMS MS#M114892.

†Computational Brain Science Lab, Department of Computational Science and Technology, School of Computer Science and Communication, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden (tony@kth.se, https://www.kth.se/profile/tony).

By differentiating the 1-D version of (18) with respect to scale s and setting $x = 0$ and $x = \omega_0 \pi/2$ respectively as well as $\omega_0 = 1$ in the resulting equation, we obtain the following algebraic equations for how the minimum and maximum scale values depend on the relative post-smoothing scale c, Γ_s and C_s for a 1-D sine wave with angular frequency ω_0:

$$C_s \left(2c^2 + 1\right) s^2 - e^{2c^2 s} \left(C_s s^2 - 2C_s s + C_s s \Gamma_s + s - 1 + \Gamma_s\right) - s \left(2C_s + 2c^2 + 1\right) + \Gamma_s (C_s s - 1) + 1 = 0, \tag{SM3}$$

$$C_s - 2C_s c^2 s^2 - e^{2c^2 s} \left(C_s s^2 - 2C_s s + C_s s \Gamma_s + s - 1 + \Gamma_s\right) - C_s s^2 + 2C_s s - C_s s \Gamma_s + 2c^2 s + s - 1 + \Gamma_s = 0. \tag{SM4}$$

By defining functions $S_{sine,1}(\Gamma_s, c, C_s)$ and $S_{sine,1}(\Gamma_s, c, C_s)$ that represent the solutions $\hat{s}_{min}$ and $\hat{s}_{max}$ of these equations as function of the parameters Γ_s, c and C_s for $\omega_0 = 1$, the minimum and maximum scale values can because of the self-similarity over scale for an arbitrary angular frequency ω_0 of the sine wave be expressed as:

$$\hat{s}_{min} = \frac{S_{sine,1}(\Gamma_s, c, C_s)}{\omega_0^2}, \qquad \hat{s}_{max} = \frac{S_{sine,2}(\Gamma_s, c, C_s)}{\omega_0^2}. \tag{SM5}$$

Table SM1 shows numerical values of these entities for different values of the scale normalization parameter Γ_s and the relative post-smoothing scale c.

Table SM1

Minimum and maximum scale estimates $\hat{s}_{min} = S_{sine,1}(\Gamma_s, c, C_s)/\omega_0^2$ and $\hat{s}_{max} = S_{sine,2}(\Gamma_s, c, C_s)/\omega_0^2$ for dense scale selection based on local extrema over scale of the post-smoothed quasi quadrature entity $\overline{Q}_{(x,y),\Gamma-norm} L$ applied to a 1-D sine wave with angular frequency $\omega_0 = 1$ for different values of the relative scale normalization parameter Γ_s and the relative post-smoothing scale c with C_s according to (25). These entities are defined as the solutions of equations (SM3) and (SM4), and for increasing values of c they approach $\hat{s}_{intermed} = \sqrt{(1 - \Gamma_s)(2 - \Gamma_s)}/\omega_0^2$ in (24).

Minimum relative scale estimate $\hat{s}_{min} = S_{sine,1}(\Gamma_s, c, C_s)$ for a 1-D sine wave

c	$\Gamma_s = 0$	$\Gamma_s = \frac{1}{4}$	$\Gamma_s = \frac{1}{2}$
0	1.000	0.750	0.500
1/2	1.033	0.741	0.466
$1/\sqrt{2}$	1.132	0.772	0.446
1	1.329	0.963	0.451
$\sqrt{2}$	1.408	1.125	0.779
2	1.414	1.145	0.863

Maximum relative scale estimate $\hat{s}_{min} = S_{sine,2}(\Gamma_s, c, C_s)$ for 1-D sine wave

c	$\Gamma_s = 0$	$\Gamma_s = \frac{1}{4}$	$\Gamma_s = \frac{1}{2}$
0	2.000	1.750	1.500
1/2	1.701	1.485	1.264
$1/\sqrt{2}$	1.584	1.363	1.147
1	1.474	1.242	1.021
$\sqrt{2}$	1.420	1.163	0.914
2	1.414	1.146	0.869

Whereas the functions $S_{sine,1}(\Gamma_s, c, C_s)$ and $S_{sine,1}(\Gamma_s, c, C_s)$ are not expressed in terms of elementary functions, it is straightforward to implement these functions using standard numerical methods for computing the solutions of a 1-D equation.

SM1.2. Phase-compensated scale estimates with post-smoothing. Given these expressions for the minimum and maximum scale estimates for a sine wave, we can also define a corresponding notion of phase compensation in the presence of spatial post-smoothing:

$$\hat{s}_{\overline{\mathcal{Q}}L,comp} = \frac{\sqrt{S_{sine,1}(\Gamma_s, c, C_s)\, S_{sine,2}(\Gamma_s, c, C_s)}}{(\overline{\mathcal{Q}}_{(x,y),1,\Gamma-norm} L + \overline{\mathcal{Q}}_{(x,y),2,\Gamma-norm} L)} \times \left(\frac{\overline{\mathcal{Q}}_{(x,y),1,\Gamma-norm} L}{S_{sine,1}(\Gamma_s, c, C_s)} + \frac{\overline{\mathcal{Q}}_{(x,y),2,\Gamma-norm} L}{S_{sine,2}(\Gamma_s, c, C_s)} \right) \hat{s}_{\overline{\mathcal{Q}}L} \tag{SM6}$$

or

$$\hat{s}_{\mathcal{Q}L,comp} = \sqrt{S_{sine,1}(\Gamma_s, c, C_s)\, S_{sine,2}(\Gamma_s, c, C_s)}\, \hat{s}_{\mathcal{Q}L} / (S_{sine,1}(\Gamma_s, c, C_s))^{\frac{\overline{\mathcal{Q}}_{(x,y),1,\Gamma-norm} L}{\overline{\mathcal{Q}}_{(x,y),1,\Gamma-norm} L + \overline{\mathcal{Q}}_{(x,y),2,\Gamma-norm} L}} / (S_{sine,2}(\Gamma_s, c, C_s))^{\frac{\overline{\mathcal{Q}}_{(x,y),2,\Gamma-norm} L}{\overline{\mathcal{Q}}_{(x,y),1,\Gamma-norm} L + \overline{\mathcal{Q}}_{(x,y),2,\Gamma-norm} L}} \tag{SM7}$$

with $\overline{\mathcal{Q}}_{(x,y),1,\Gamma-norm} L$ and $\overline{\mathcal{Q}}_{(x,y),2,\Gamma-norm} L$ according to (12) and (13) and with the normalization chosen such that the scale values should aim towards the geometric mean of the extreme values $S_{sine,1}(\Gamma_s, c, C_s)$ and $S_{sine,2}(\Gamma_s, c, C_s)$ according to (SM3) and (SM4).

SM1.2.1. Two-dimensional blob. For a two-dimensional Gaussian blob

$$f(x, y) = g(x, y;\ s_0) = \frac{1}{2\pi s_0} e^{-(x^2+y^2)/2s_0}, \tag{SM8}$$

if follows from the semi-group property of the Gaussian that the scale-space representation is given by

$$L(x, y;\ s) = g(x, y;\ s_0 + s) = \frac{1}{2\pi(s_0 + s)} e^{-(x^2+y^2)/2(s_0+s)} \tag{SM9}$$

and the unsmoothed quasi quadrature entity at the origin is of the form

$$\mathcal{Q}_{(x,y),\Gamma-norm} L = \frac{C_s s^{2-\Gamma_s}}{2\pi^2 (s + s_0)^4}. \tag{SM10}$$

Differentiating this expression with respect to the scale parameter s, shows that the maximum value over scales is assumed at scale

$$\hat{s}_{Gauss} = \frac{2 - \Gamma_s}{2 + \Gamma_s}\, s_0. \tag{SM11}$$

With complementary post-smoothing with relative integration scale c, corresponding computation of the post-smoothed quasi quadrature and differentiation of the resulting expression gives an algebraic equation of the form

$$\begin{aligned}
&4c^6 s^2 \left((2C_s+1)s^2(2+\Gamma_s) \; +2ss_0(C_s(\Gamma_s-1)+1+\Gamma_s)+s_0^2\Gamma_s\right)+\\
&+4c^4 s(s+s_0)\left((2C_s+1)s^2(2+\Gamma_s) \; +ss_0(2C_s(\Gamma_s-2)+1+2\Gamma_s)+s_0^2(\Gamma_s-1)\right)+\\
&+c^2(s+s_0)^2\left((4C_s+1)s^2(2+\Gamma_s) \; +2ss_0(2C_s(\Gamma_s-1)+\Gamma_s)+s_0^2(\Gamma_s-2)\right)\\
&+C_s(s+s_0)^3(s(2+\Gamma_s)+s_0(\Gamma_s-2))=0
\end{aligned} \tag{SM12}$$

for the selected scale level $\hat{s}$ as function of the complementary scale normalization parameter Γ_s, the relative post-smoothing scale c and the relative weighting factor C_s between first- and second-order information. Let us define the following function for representing the solution of this equation:

$$\hat{s} = S_{Gauss}(\Gamma_s, c, C_s)\, s_0. \tag{SM13}$$

Table SM2 shows numerical values of this entity for different values of Γ_s and c.

Table SM2

Numerical values of the ratio $\hat{s}/s_0 = S_{Gauss}(\Gamma_s, c, C_s)$ for which the post-smoothed quasi quadrature entity $\overline{\mathcal{Q}}_{(x,y),\Gamma-norm}L$ assumes its maximum over scale at the center of a Gaussian blob with scale parameter s_0 for different values of the complementary scale normalization parameter Γ_s and the relative post-smoothing scale parameter c with C_s according to (25).

Scale estimates $\hat{s} = S_{Gauss}(\Gamma_s, c)\, s_0$ based on $\overline{\mathcal{Q}}_{(x,y),\Gamma-norm}L$ at the center of a Gaussian blob

c	$\Gamma_s = 0$	$\Gamma_s = \frac{1}{4}$	$\Gamma_s = \frac{1}{2}$
0	1.000	0.778	0.600
1/2	0.839	0.650	0.498
$1/\sqrt{2}$	0.751	0.578	0.440
1	0.641	0.487	0.367
$\sqrt{2}$	0.519	0.385	0.283
2	0.402	0.285	0.199

The difference in net effect between the Gaussian scale calibration model and the sine wave scale calibration model under variations of Γ_s and c is essentially determined by the variation of the following ratio between the scale calibration factors in units of $\sigma_s = \sqrt{s}$:

$$\chi(\Gamma_s, c) = \sqrt{\frac{2S_{Gauss}(\Gamma_s, c)}{S_{sine,2}(\Gamma_s, c)}} \tag{SM14}$$

see Table SM3 for numerical values. For $c \leq 1$, it can be seen that the relative differences in effects for scale calibration are within a range of 15 % in units of $\sigma_s = \sqrt{s}$.

Given the similarity between the results obtained from these qualitatively very different models, it seems plausible that the results should also generalize to wider classes of image structures. Choosing the parameter γ for scale selection using γ-normalized derivatives based on the behaviour for Gaussian image models has also been demonstrated to lead to highly useful results for a wide range of computer vision tasks (Lindeberg [SM2, SM1, SM3, SM4]).

Table SM3

Numerical values of the ratio between the scale calibration factors $S_{Gauss}(\Gamma_s, c)$ and $S_{sine,2}(\Gamma_s, c)$ in units of $\sigma_s = \sqrt{s}$ and normalized such that the ratio is equal to one for $\Gamma_s = 0$ and $c = 0$. This ratio provides an estimate of how much the calibrated scale estimates will depend on the choice of calibration model, based on either sparse image features or a dense texture pattern.

Dependency of the ratio between the scale calibration factors $\sqrt{\frac{2S_{Gauss}(\Gamma_s,c)}{S_{sine,2}(\Gamma_s,c)}}$ on Γ_s and c

c	$\Gamma_s = 0$	$\Gamma_s = \frac{1}{4}$	$\Gamma_s = \frac{1}{2}$
0	1.000	0.943	0.894
1/2	0.991	0.935	0.888
$1/\sqrt{2}$	0.974	0.921	0.876
1	0.933	0.886	0.848
$\sqrt{2}$	0.855	0.814	0.786
2	0.754	0.705	0.677

SM2. Detailed analysis of phase compensation and scale calibration for dense spatio-temporal scale selection. For dense spatio-temporal scale selection, we do according to Section 4 at every point (x, y, t) in space-time detect simultaneous local extrema over spatio-temporal scales (81)

$$(\hat{s}_{\mathcal{Q}_{(x,y,t),\Gamma-norm}}, \hat{\tau}_{\mathcal{Q}_{(x,y,t),\Gamma-norm}}) = \text{argmaxlocal}_{s,\tau}\, \mathcal{Q}_{(x,y,t),\Gamma-norm} L \tag{SM15}$$

of the scale-normalized spatio-temporal quasi quadrature entity (73)

$$\begin{aligned}
&\mathcal{Q}_{(x,y,t),\Gamma-norm} L \\
&= \frac{\tau\, \mathcal{Q}_{(x,y),\Gamma-norm} L_t + C_\tau \tau^2\, \mathcal{Q}_{(x,y),\Gamma-norm} L_{tt}}{\tau^{\Gamma_\tau}} \\
&= \frac{1}{s^{\Gamma_s} \tau^{\Gamma_\tau}} \left(\tau \left(s \left(L_{xt}^2 + L_{yt}^2 \right) + C_s\, s^2 \left(L_{xxt}^2 + 2L_{xyt}^2 + L_{yyt}^2 \right) \right) \right. \\
&\qquad \left. + C_\tau\, \tau^2 \left(s \left(L_{xtt}^2 + L_{ytt}^2 \right) + C_s\, s^2 (L_{xxtt}^2 + 2L_{xytt}^2 + L_{yytt}^2) \right) \right) .
\end{aligned} \tag{SM16}$$

SM2.1. Phase-compensated scale estimates. Given the understanding from Section 4.3 of how the local spatio-temporal scale estimates depend on the local phase of a sine wave, we can define *phase-compensated spatial and temporal scale estimates* according to

$$\hat{s}_{\mathcal{Q}_{(x,y,t)L,comp}} = \tag{SM17}$$

$$\frac{\sqrt{(1-\Gamma_s)(2-\Gamma_s)}\, \hat{s}_{\mathcal{Q}_{(x,y,t)L,\Gamma-norm}}}{(1-\Gamma_s) \frac{\mathcal{Q}_{1,(x,y),\Gamma-norm} L_t + C_\tau \mathcal{Q}_{1,(x,y),\Gamma-norm} L_{tt}}{\mathcal{Q}_{(x,y,t),\Gamma-norm} L} (2-\Gamma_s) \frac{\mathcal{Q}_{2,(x,y),\Gamma-norm} L_t + C_\tau \mathcal{Q}_{2,(x,y),\Gamma-norm} L_{tt}}{\mathcal{Q}_{(x,y,t),\Gamma-norm} L}},$$

$$\hat{\tau}_{\mathcal{Q}_{(x,y,t)L,comp}} = \tag{SM18}$$

$$\frac{\sqrt{(1-\Gamma_\tau)(2-\Gamma_\tau)}\, \hat{\tau}_{\mathcal{Q}_{(x,y,t)L,\Gamma-norm}}}{(1-\Gamma_\tau) \frac{\mathcal{Q}_{1,(x,y),\Gamma-norm} L_t + \mathcal{Q}_{2,(x,y),\Gamma-norm} L_t}{\mathcal{Q}_{(x,y,t),\Gamma-norm} L} (2-\Gamma_\tau) \frac{C_\tau \left(\mathcal{Q}_{1,(x,y),\Gamma-norm} L_{tt} + \mathcal{Q}_{2,(x,y),\Gamma-norm} L_{tt} \right)}{\mathcal{Q}_{(x,y,t),\Gamma-norm} L}},$$

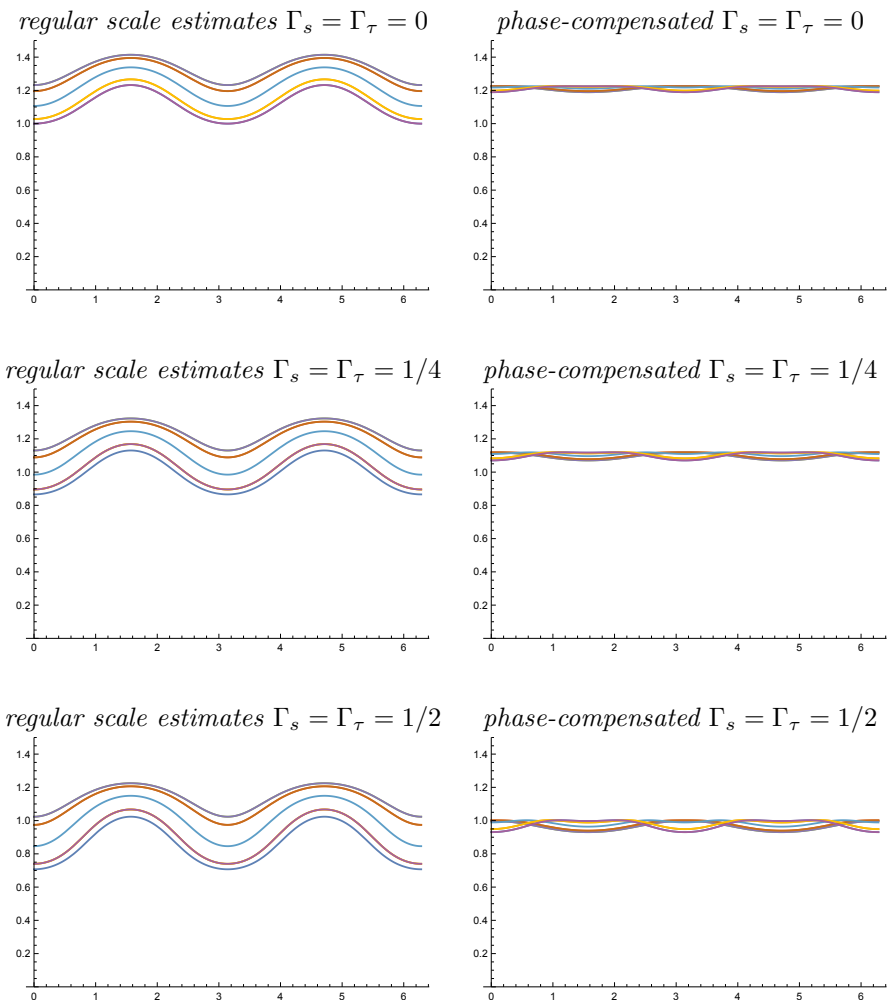

Figure SM1. *Spatial variability of* spatial scale estimates $\hat{\sigma}_s = \sqrt{s}$ *computed (left column) from local extrema over spatial scale of the* Γ*-normalized quasi quadrature measure* $\mathcal{Q}_{(x,y,t),\Gamma-norm} L$ *according to* (81) *for a* $2+1$*-D spatio-temporal sine wave with spatial angular frequency* $\omega_s = 1$ *and (right column) with phase compensation of the scale estimates according to* (SM17)*. (Horizontal axis: spatial position* t*) (The multiple graphs in each diagram show the variability of the scale estimates for different values of the complementary spatial coordinate* $y = n\pi/8$*.)*

defined to be equal to the geometric averages of the extreme values

$$\hat{s} = \sqrt{\hat{s}_1 \, \hat{s}_2} = \frac{\sqrt{(1-\Gamma_s)(2-\Gamma_s)}}{\omega_0^2} \tag{SM19}$$

$$\hat{\tau} = \sqrt{\hat{\tau}_1 \, \hat{\tau}_2} = \frac{\sqrt{(1-\Gamma_\tau)(2-\Gamma_\tau)}}{\omega_\tau^2} \tag{SM20}$$

in the extreme cases when only one of the first- or second-order components in the components responds and with a much lower variability in between because of the blending of the responses

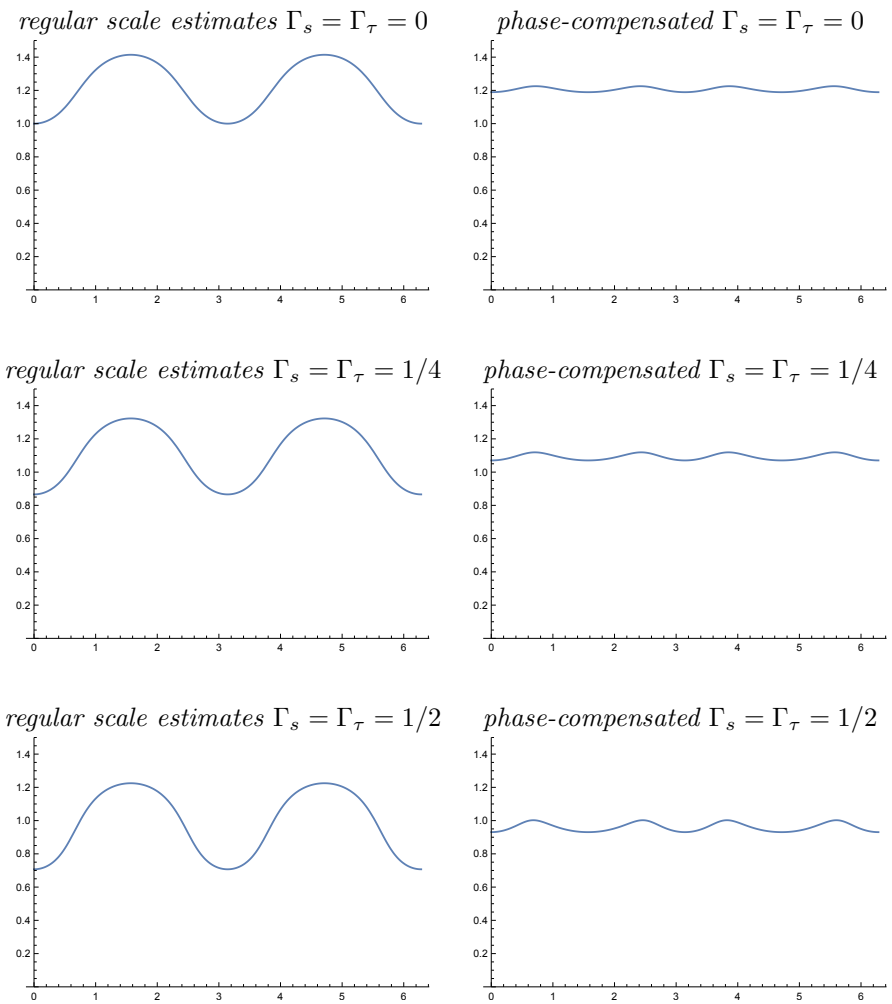

Figure SM2. *Temporal variability of* temporal scale estimates $\hat{\sigma}_\tau = \sqrt{\tau}$ *computed (left column) from local extrema over temporal scale of the* Γ*-normalized quasi quadrature measure* $\mathcal{Q}_{(x,y,t),\Gamma-norm} L$ *according to* (81) *for a* $2+1$*-D spatio-temporal sine wave with spatial angular frequency* $\omega_s = 1$ *and (right column) with phase compensation of the scale estimates according to* (SM18)*. (Horizontal axis: time* t*)*

to the first- *vs.* second-order spatial or temporal derivatives (see figures SM1–SM2).

From these spatial and temporal scale estimates, we can in turn estimate the spatial and temporal wavelengths of the sine wave according to

$$\hat{\lambda}_s = \frac{2\pi \sqrt{\hat{s}_{\mathcal{Q}_{(x,y,t)L,comp}}}}{\sqrt[4]{(1-\Gamma_s)(2-\Gamma_s)}}, \tag{SM21}$$

$$\hat{\lambda}_\tau = \frac{2\pi \sqrt{\hat{\tau}_{\mathcal{Q}_{(x,y,t)L,comp}}}}{\sqrt[4]{(1-\Gamma_\tau)(2-\Gamma_\tau)}}. \tag{SM22}$$

SM2.2. Scale calibration. When applied to a Gaussian blink of spatial extent s_0 and temporal duration τ_0

$$f(x,y,t) = g(x,y;\ s_0)\, g(t;\ \tau_0) = \frac{1}{(2\pi)^{3/2} s_0 \sqrt{\tau_0}}\, e^{-(x^2+y^2)/2s_0}\, e^{-t^2/2\tau_0}, \tag{SM23}$$

for which the spatio-temporal scale-space representation is of the form

$$L(x,y,t;\ s_0,\tau_0) = g(x,y;\ s_0+s)\, g(t;\ \tau_0+\tau), \tag{SM24}$$

the spatial and temporal scale estimates will according to the theoretical analysis in (Lindeberg [SM5, SM6]) be given by

$$\hat{s} = \frac{\gamma_s s_0}{2-\gamma_s}\, s_0, \tag{SM25}$$

$$\hat{\tau} = \frac{2\gamma_\tau \tau_0}{3-2\gamma_\tau}\, \tau_0. \tag{SM26}$$

Here, for $\gamma_s = 1 - \Gamma_s/2$ and $\gamma_\tau = 1 - \Gamma_\tau/2$ this implies that the regular scale estimates for a Gaussian blink are

$$\hat{s} = \frac{2-\Gamma_s}{2+\Gamma_s}\, s_0, \tag{SM27}$$

$$\hat{\tau} = \frac{2-\Gamma_\tau}{1+\Gamma_\tau}\, \tau_0, \tag{SM28}$$

and the corresponding phase-compensated scale estimates

$$\hat{s}_{\mathcal{Q}_{(x,y,t)L,comp}} = \frac{\sqrt{(2-\Gamma_s)(1-\Gamma_s)}}{2+\Gamma_s}\, s_0, \tag{SM29}$$

$$\hat{\tau}_{\mathcal{Q}_{(x,y,t)L,comp}} = \frac{\sqrt{(2-\Gamma)(1-\Gamma)}}{1+\Gamma_\tau}\, \tau_0. \tag{SM30}$$

If we want to calibrate the spatial and temporal scale estimates such that the spatial and temporal scale estimates are equal to $\hat{s} = s_0$ and $\hat{\tau} = \tau_0$ for a Gaussian blink of spatial extent s_0 and temporal duration τ_0, we should therefore calibrate the phase compensated scale estimates $\hat{s}_{\mathcal{Q}_{(x,y,t)L,comp}}$ and $\hat{\tau}_{\mathcal{Q}_{(x,y,t)L,comp}}$ according to

$$\hat{s}_{\mathcal{Q}_{(x,y,t)L,calib}} = \frac{2+\Gamma_s}{\sqrt{(2-\Gamma_s)(1-\Gamma_s)}}\, \hat{s}_{\mathcal{Q}_{(x,y,t)L,comp}}, \tag{SM31}$$

$$\hat{\tau}_{\mathcal{Q}_{(x,y,t)L,calib}} = \frac{1+\Gamma_\tau}{\sqrt{(2-\Gamma_\tau)(1-\Gamma_\tau)}}\, \hat{\tau}_{\mathcal{Q}_{(x,y,t)L,comp}}. \tag{SM32}$$

With this scale calibration, since the scale estimate for a Gaussian temporal onset ramp, which for regular γ-normalized temporal derivatives assumes the form (Lindeberg [SM1, equation (23)])

$$\hat{s} = \frac{\gamma}{1-\gamma}\, s_0 = \{\gamma = 1-\Gamma\} = \frac{1-\Gamma}{\Gamma}\, s_0, \tag{SM33}$$

the spatial scale estimate for a diffuse Gaussian edge will by combination of (SM17) with (SM31) be given by

$$\hat{s} = \frac{2 + \Gamma_s}{\Gamma_s} \, s_0, \tag{SM34}$$

whereas the temporal scale estimate for a Gaussian onset ramp will by combination of (SM18) with (SM32) be given by

$$\hat{\tau} = \frac{1 + \Gamma_\tau}{\Gamma_\tau} \, \tau_0. \tag{SM35}$$

By varying the parameters Γ_s and Γ_τ we can thereby regulate the factor by which the spatial scale estimate for a diffuse Gaussian edge will be proportional to its diffuseness and in a corresponding manner the factor by which the temporal scale estimate for a Gaussian onset ramp will be proportional to its temporal duration, while ensuring that the spatial and temporal scale estimates for a Gaussian blink will still reflect the spatial extent and the temporal duration of the Gaussian blink.

SM2.3. Phase-compensated magnitude estimates. When performing temporal scale selection from the local extrema of the scale-normalized quasi quadrature measure over scale (73), the magnitude responses at the spatial points $(x = n\pi/\omega_s, y = n\pi/\omega_s)$ and the temporal moments $t = n\pi/\omega_\tau$ at which only the first-order temporal derivative responds are given by

$$\mathcal{Q}_{t,1,\Gamma-norm} = 2(1 - \Gamma_s)^{1-\Gamma_s} (1 - \Gamma_\tau)^{1-\Gamma_\tau} \, e^{\Gamma_s + \Gamma_\tau - 2} \, \omega_s^{2\Gamma_s} \, \omega_\tau^{2\Gamma_\tau}, \tag{SM36}$$

whereas the magnitude responses at the spatial points $(x = (\pi/2 + n\pi)/\omega_s, y = (\pi/2 + n\pi)/\omega_s)$ and the temporal moments $t = (\pi/2 + n\pi)/\omega_0^2$ at which only the second-order temporal derivative responds are given by

$$\mathcal{Q}_{t,2,\Gamma-norm} = \frac{2(2 - \Gamma_s)^{2-\Gamma_s} (2 - \Gamma_\tau)^{2-\Gamma_\tau} \, e^{\Gamma_s + \Gamma_\tau - 4} \, \omega_s^{2\Gamma_s} \, \omega_\tau^{2\Gamma_\tau}}{\sqrt{(2 - \Gamma_s)(1 - \Gamma_s)} \sqrt{(2 - \Gamma_\tau)(1 - \Gamma_\tau)}}. \tag{SM37}$$

Thus, the magnitude responses will have a certain phase dependency because of the variability in the temporal scale estimates leading to corresponding variability in the relative strengths of the first- *vs.* second-order responses (see the left columns in figures SM3–SM4). When performing phase compensation according to (SM17) and (SM18), the temporal scale estimates will on the other hand will be close to a temporal scale level where the relative strengths of the first- and second-order responses are balanced and leading to a much lower temporal variability in the magnitude responses (see the right columns in figures SM3–SM4). If one additionally wants these magnitude estimates to be independent of the wavelengths of the sine wave pattern, then this can be accomplished by instead computing the corresponding post-normalized quasi quadrature entity

$$\mathcal{Q}_{(x,y,t),post-norm} = \hat{s}^{\Gamma_s} \, \hat{\tau}^{\Gamma_\tau} \, \mathcal{Q}_{(x,y,t),\Gamma-norm} \tag{SM38}$$

where $\hat{s}$ and $\hat{\tau}$ represent the phase-compensated spatial and temporal scale estimates according to (SM17) and (SM18).

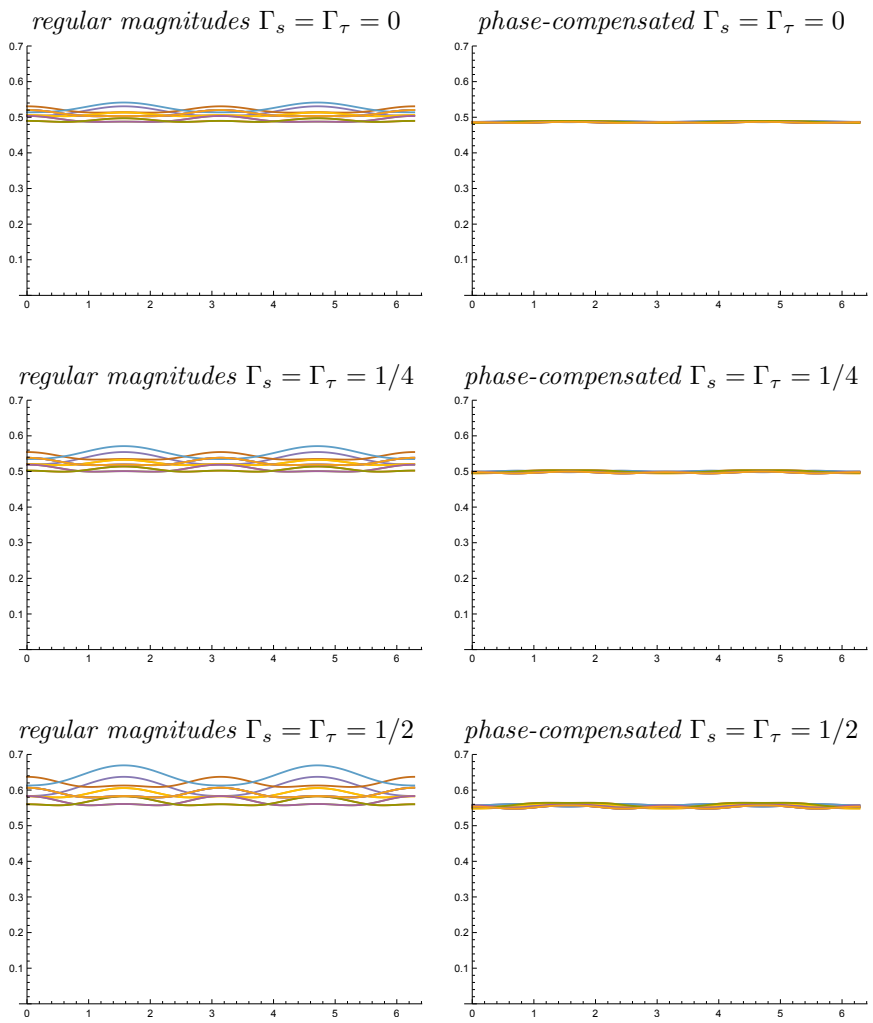

Figure SM3. *Spatial variability of magnitude responses $\mathcal{Q}_{(x,y,t),\Gamma-norm}$ computed (left column) at the local extrema over both spatial and temporal scales of the Γ-normalized quasi quadrature measure $\mathcal{Q}_{(x,y,t),\Gamma-norm}$ according to (81) for a $2+1$-D spatio-temporal sine wave pattern of spatial angular frequency $\omega_s = 1$ and temporal angular frequency $\omega_\tau = 1$ (right column) at phase-compensated scale estimates according to (SM17) and (SM18). (Horizontal axis: spatial coordinate x) (The multiple graphs in each diagram show the variability of the scale estimates for different values of the complementary spatial coordinate $y = m\pi/4$ and the temporal coordinate $t = n\pi/4$.)*

In these respects, the analysis in this supplement shows how the notion of phase compensation also applies in a spatio-temporal setting with independent variabilities in the spatial and the temporal scales in the spatio-temporal image structures in video data.

When reduced to either a purely spatial or a purely temporal domain, the analysis in this supplement also gives a more detailed treatment of how the notion of scale calibration can be performed when applying dense scale selection to either purely spatial image data or a purely temporal signal. Specifically, the expressions (SM34) and (SM35) show how variations

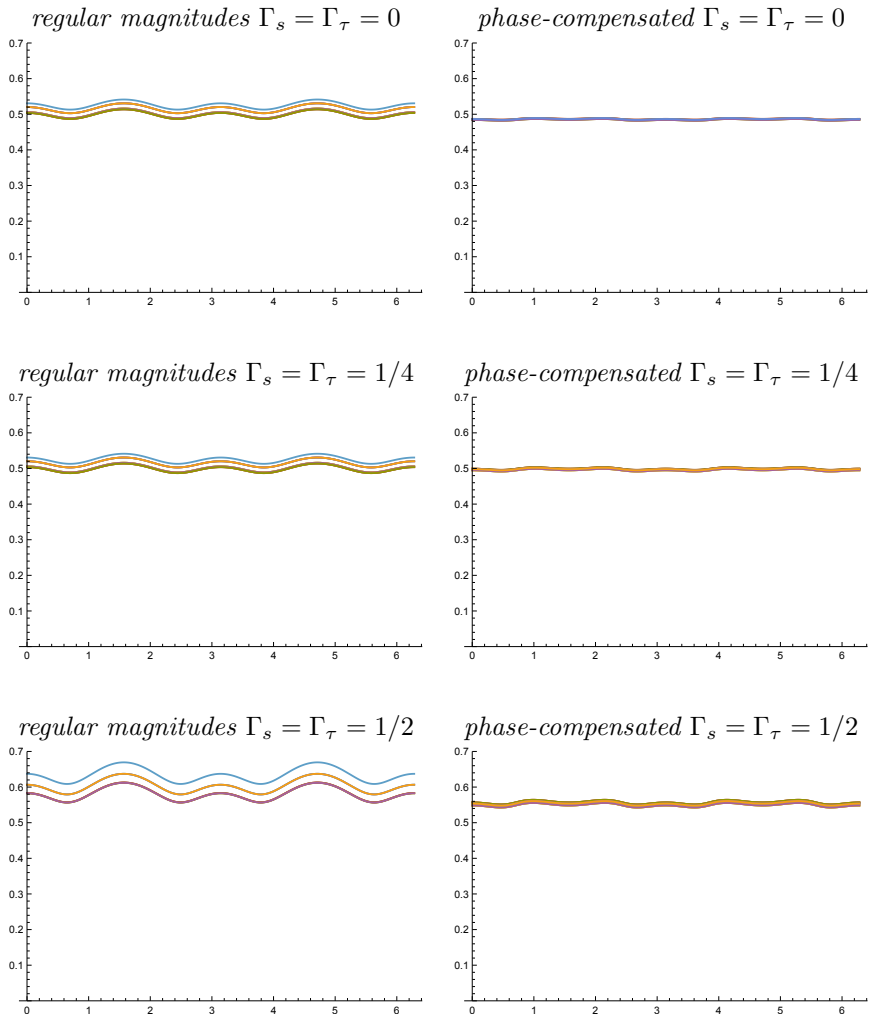

Figure SM4. *Temporal variability of magnitude responses $\mathcal{Q}_{(x,y,t),\Gamma-norm}$ computed (left column) at the local extrema over both spatial and temporal scales of the Γ-normalized quasi quadrature measure $\mathcal{Q}_{(x,y,t),\Gamma-norm}$ according to (81) for a $2+1$-D spatio-temporal sine wave pattern of spatial angular frequency $\omega_s = 1$ and temporal angular frequency $\omega_\tau = 1$ (right column) at phase-compensated scale estimates according to (SM17) and (SM18). (Horizontal axis: time t) (The multiple graphs in each diagram show the variability of the scale estimates for different values of the complementary spatial coordinates $x = m\pi/4$ and $y = n\pi/4$.)*

in the complementary scale normalization parameters Γ_s and Γ_τ will influence the selection of spatial and temporal scales at diffuse spatial edges and temporal ramps.

REFERENCES

[1] T. Lindeberg, *Edge detection and ridge detection with automatic scale selection*, International Journal of Computer Vision, 30 (1998), pp. 117–154.

[2] T. Lindeberg, *Feature detection with automatic scale selection*, International Journal of Computer Vision, 30 (1998), pp. 77–116.

[3] T. Lindeberg, *Principles for automatic scale selection*, in Handbook on Computer Vision and Applications, Academic Press, Boston, USA, 1999, pp. 239–274. Also available from http://www.csc.kth.se/cvap/abstracts/cvap222.html.

[4] T. Lindeberg, *Scale selection*, in Computer Vision: A Reference Guide, K. Ikeuchi, ed., Springer, 2014, pp. 701–713.

[5] T. Lindeberg, *Spatio-temporal scale selection in video data*, in Proc. Scale Space and Variational Methods in Computer Vision (SSVM 2017), vol. 10302 of Springer LNCS, 2017, pp. 3–15.

[6] T. Lindeberg, *Spatio-temporal scale selection in video data*, Journal of Mathematical Imaging and Vision, 2017, pp. 1–38, https://doi.org/10.1007/s10851-017-0766-9.