Temporal Scale Spaces

Daniel Fagerström

Computational Vision and Active Perception Laboratory (CVAP)
Department of Numerical Analysis and Computing Science
KTH (Royal Institute of Technology) SE-100 44 Stockholm, Sweden
danielf@nada.kth.se

Abstract. In this paper we discuss how to define a scale space suitable for temporal measurements. We argue that such a temporal scale space should possess the properties of: temporal causality, linearity, continuity, positivity, recursitivity as well as translational and scaling covariance. It is shown that these requirements imply a one parameter family of convolution kernels. Furthermore it is shown that these measurements can be realized in a time recursive way, with the current data as input and the temporal scale space as state, i.e. there is no need for storing earlier input. This family of measurement processes contains the diffusion equation on the half line (that represents the temporal scale) with the input signal as boundary condition on the temporal axis. The diffusion equation is unique among the measurement processes in the sense that it is preserves positivity (in the scale domain) and is infinitesimally generated. A numerical scheme is developed and relations to other approaches are discussed.

1 Introduction

An important difference between spatial and temporal observation is that while all spatial directions basically can be treated in the same way there is certainly a difference between moving forward and backward in time. We have no access to future observations while we could have memorized earlier observations. A temporal observer must respect *causality*.

Furthermore, if we consider a causal scale space as an idealized device for performing real time measurement of a temporal signal *the observer cannot access the past*, only its memory of the past. An important property of a theory about real time temporal measurements is therefore that it has a time recursive formulation.

There have been several proposals on how to define a *causal scale space*. Koenderink [5] and Florack [3] map the half-line between current moment and the infinite past to a line by a logarithmic transformation. They then apply Gaussian scale space on the transformed signal. Lindeberg and Fagerström [7] base their axiomatization on the non-creation of local extrema. The axiomatization leads to scale spaces where either time or scale must be discrete and their scale spaces are not scale covariant. Salden et al. [11] derive the diffusion equation from conservation principles. They adapt their theory to the temporal domain by applying the diffusion equation on the past half-line, and by imposing reflecting boundary conditions.

The causal scale space theories of Koenderink [5], Florack [3] and Salden et al. [11] are formulated in terms of convolution against or diffusion on the past signal and have to our knowledge, no time recursive realization. The causal scale space theory of Lindeberg and Fagerström [7], has a time recursive formulation, but its lack of continuous formulation and scale covariance makes it less convenient to use as a basic theory about temporal observation.

In this article we will require a causal scale space to be temporally causal, linear, continuous, positive, having translational and scaling covariance and having a semigroup property. These requirements are the same as Pauwels et al. [8] used for defining scale spaces appropriate for spatial measurements, except for that we require temporal causality instead of reflectional symmetry. Each scale in the causal scale space is generated by convolving the input signal with a causal scale space kernel of a certain dilation. We show that there is a one parameter family of convolution kernels, know from probability theory as *extremal stable density functions*, that fulfills our requirements. We continue by showing that there is a time recursive realization of the causal scale spaces, using only the scale space it self as memory of earlier input. For only one of the parameter values, the recursive realization is infinitesimally generated. For this parameter the temporal scale space is given by the diffusion equation on the half line, a numerical scheme is given for this special case. We conclude by a comparison with earlier works.

2 Temporal Measurement

An temporal signal is usually thought about as a function from time coordinates to scalar values, $u(t)$. However, measurement of the instantaneous value of a signal is physically impossible, each measurement must take a non vanishing amount of time. Following the approach advocated by Florack [3] we instead start by designing our measurement apparatus, our space of *test functions* $\phi \in \Delta$, and try to make them as point like as possible . We then consider the temporal signal as a "black box", $u \in \Sigma$, that we can probe with our test functions, $u(\phi)$. Even if we cannot measure the instantaneous value of the signal we still need to associate each measurement with a certain moment, so that the measurement can be said to be performed at time t. To accomplish this a *measurement time*:

$$\pi(\phi) = t, \qquad \pi : \Delta \to \mathbb{R},$$

is defined for each test function, and the notation $\phi_t \in \pi^{-1}(t)$, is used for denoting a test function applied at time t.

For a seeing system with a wide range of visual competences it seems reasonable that the representation of the visual measurement as far as possible should avoid semantic interpretation of the environment, it should be *uncommitted* [6]. The representation of the input should just *embody* general structure in the environment.

2.1 Linearity

The measurement process is supposed to be linear,

$$u(\alpha\phi + \beta\psi) = \alpha u(\phi) + \beta u(\psi),$$

with $\alpha, \beta \in \mathrm{I\!R}$. We also require that the sensor functions are well behaved in the following sense:

$$\Delta \subset L_1 \cap C^\infty \cap \{\phi | \phi(-\infty) = \phi(\infty) = 0\}.$$

The signal space Σ is defined as the topologically dual space to the space of test functions, $\Sigma = \Delta'$ [14]. In many cases a signal $u \in \Sigma$ can be represented, using the Riesz representation theorem, as $u(\phi) = \int u'(t)\phi(t)\,dt$, for some function u'. Differentiation in the signal space can be defined by

$$(\partial_t u)(\phi) = -u(\partial_t \phi). \tag{1}$$

We don't have to worry about the differentiability of the signals, it suffices to have differentiable test functions. The action of a diffeomorphism f on a signal can be defined as:

$$fu(\phi) = u(|J_{f^{-1}}|\phi \circ f^{-1}). \tag{2}$$

Our main task is to define a *minimal* space of test functions, that suits the needs of the observation task.

2.2 Causality

A temporal measurement should not involve any future information.

$$\phi_t(t') = 0, \forall t' > t \tag{3}$$

This formalizes the differences between time and space that was discussed above.

2.3 Covariance

The structure of the sensor system should correspond to regularities in the environment. There is no a priori reason to believe that a certain moment of time, or a certain span of time should posses properties different from all the rest. This should be reflected in the measurement process: i.e. it should be *translation covariant*, the measurement process should be the same for each moment of time, and it should be *scaling covariant*, all length of time spans should be treated in the same way. To obtain translation and scaling covariance, a family of test functions $\phi'_{t,\tau} : \mathrm{I\!R}\mathrm{I\!R}_+ \to \Delta$, indexed by time t and scale τ is needed. Measurement at scale τ, is denoted by:

$$\Phi_\tau u(t) = u(\phi'_{t,\tau}), \qquad \Phi_\tau : \Sigma \to C^\infty. \tag{4}$$

Measurement is translation covariant if

$$\Phi_\tau T_a = T_a \Phi_\tau, \tag{5}$$

where $T_a f(x) = f(x-a)$. From the left hand side we get,

$$\Phi_\tau T_a u(t) = T_a u(\phi'_{t,\tau}) = u(T_{-a}\phi'_{t,\tau}),$$

by using equation (2), and from the right hand side we get,

$$T_a \Phi_\tau u(t) = \Phi_\tau u(t-a) = u(\phi'_{t-a,\tau})$$

As this holds for all u we get,

$$T_a \phi'_{t,\tau} = \phi'_{t+a,\tau},$$

and

$$\phi'_{t,\tau}(t') = T_t \phi'_{0,\tau}(t') = \phi'_{0,\tau}(t'-t) = \phi_{0,\tau}(t-t'),$$

where $\phi_{t,\tau}(t) = \phi'_{t,\tau}(-t)$. Using this in equation (4), we get,

$$\Phi_\tau u(t) = u(\phi_{0,\tau}(t-\cdot)) = u * \phi_{0,\tau}(t) = u * \phi_\tau(t), \tag{6}$$

where the convolution is in distribution sense and $\phi_\tau = \phi_{0,\tau}$. We also note that while convolution kernels are reflected compared to measurement functions, the temporal causality requirement for convolution kernels becomes:

$$\phi_\tau(t) = 0, \forall t < 0. \tag{7}$$

Scaling covariance has a somewhat more complicated form:

$$\Phi_\tau S_\gamma = S_\gamma \Phi_{\psi(\gamma)\tau}, \tag{8}$$

where $S_\gamma f(x) = f(\gamma x)$ and ψ is an invertible function such that $\psi(0) = 0$. Scaling up a signal and then applying measurement devices of a certain size corresponds to using smaller measurement devices at the original signal and performing the scaling afterward. Why there is a need for a coordinate change ψ on the rescaling parameter γ, will be clear later. From the left hand side of equation (8), and equation (2) we get

$$\Phi_\tau S_\gamma u(t) = S_\gamma u * \phi_\tau(t) = u(D_\gamma \phi_\tau(t-\cdot)),$$

where $D_\gamma f(x) = (1/\gamma) f(x/\gamma)$, and from the right hand side we get,

$$S_\gamma \Phi_{\psi(\gamma)\tau} u(t) = \Phi_{\psi(\gamma)\tau} u(\gamma t) = u(\phi_{\psi(\gamma)\tau}(\gamma t - \cdot)).$$

Combining these two expressions and setting $t = 0$ we have,

$$D_\gamma \phi_\tau = \phi_{\psi(\gamma)\tau}. \tag{9}$$

2.4 Cascade Property

The result of a measurement of a signal can be considered to be a signal in turn. It then seems reasonable that a measurement of a measurement should correspond to single measurement. This can be formalized as:

$$\Phi_\tau \Phi_{\tau'} = \Phi_{\tau+\tau'}, \tag{10}$$

that is measurements form a semigroup. The semigroup property of measurement means that the measurement kernels form a convolution algebra. From the left hand side,

$$\Phi_\tau \Phi_{\tau'} u = (\Phi_{\tau'} u) * \phi_\tau = (u * \phi_{\tau'}) * \phi_\tau = u * (\phi_{\tau'} * \phi_\tau),$$

and from the right hand side, $\Phi_{\tau+\tau'} u = u * \phi_{\tau+\tau'}$, and by combining these we get,

$$\phi_\tau * \phi_{\tau'} = \phi_{\tau+\tau'}. \tag{11}$$

2.5 Extended Point

As mentioned earlier a measurement should approach a pointwise value of the signal. This could be described as:

$$\lim_{\tau \to 0} \phi_\tau = \delta, \tag{12}$$

the measurement kernel approaches a Dirac pulse. We also want the measurement kernel to have *unit area*:

$$\int \phi_\tau = 1, \tag{13}$$

and to be *positive*,

$$u \geq 0 \Rightarrow u(\phi_\tau) \geq 0. \tag{14}$$

3 Characterization of causal scale space kernels

We can now summarize the requirements on measurement kernels for a causal scale space.

Definition 1. *$\phi_\tau(t)$ is called a causal scale space kernel if it possesses the properties of: Continuity, Positivity (Eq. 14), Unit area (Eq. 13), Temporal causality (Eq. 7), Dilation covariance (Eq. 9), Convolution semigroup (Eq. 11)*

Pauwels et al [8] used the same axioms, with temporal causality replaced by reflection symmetry, to characterize spatial scale spaces (see Weickert et al. [15], for comparison between different scale space axiomatizations).

We now derive the form of causal scale space kernels in a sequence of lemmas.

Lemma 1 (Convolution semigroup). *Let $\phi_\tau(t)$ be an absolutely integrable causal function, then it is a convolution semigroup iff its Laplace transform $\tilde{\phi}_\tau(s) = e^{-g(s)\tau}$, where $g(s) \in \mathbb{R}$ for $s \in \mathbb{R}$.*

Proof. From absolute integrability it can be conclude that the Laplace transform of ϕ_τ exists and is given by

$$\tilde{\phi}_\tau(s) = \mathcal{L}[\phi_\tau(\cdot)] = \int_0^\infty e^{-st}\phi_\tau(t)\,dt.$$

In the Laplace transform domain the convolution semigroup property becomes $\tilde{\phi}_{\tau_1}\tilde{\phi}_{\tau_2} = \tilde{\phi}_{\tau_1+\tau_2}$. This is an instance of Cauchy's functional equation $f(x+y) = f(x)f(y)$. For continuous real functions it has the unique solution $f(x) = e^{-cx}$, $c \in \mathbb{R}$, (see e.g. [1]) and hence

$$\tilde{\phi}_\tau(s) = e^{-g(s)\tau}, \tag{15}$$

where $g(s) \in \mathbb{R}$.

Now we will use the dilation covariance to further restrict the form of ϕ_τ.

Lemma 2 (Dilation covariant convolution semigroup). *Let $\phi_\tau(t)$ be an absolutely integrable causal function, then it is a dilation covariant convolution semigroup iff its Laplace transform $\tilde{\phi}_\tau(s) = \tilde{\phi}_{\alpha,\tau}(s) = e^{-s^\alpha\tau}$, for $s \geq 0$ and $\alpha > 0$.*

Proof. The Laplace transform of the dilation covariant equation is

$$\tilde{\phi}_\tau(\gamma s) = \tilde{\phi}_{\psi(\gamma)\tau}(s). \tag{16}$$

Substituting (15) in (16) one obtains

$$e^{-g(\gamma s)\tau} = e^{-\psi(\gamma)g(s)\tau} \tag{17}$$

and, since the exponential function is invertible,

$$g(\gamma s) = \psi(\gamma)g(s) \tag{18}$$

must hold. Without loss of generality one can assume that $g(1) = 1$, and by inserting $s = 1$ in (18), we get $\psi(\gamma) = g(\gamma)$, and (18) becomes

$$g(\gamma s) = g(\gamma)g(s). \tag{19}$$

By setting $g(t) = f(\log(t))$, we can see that (19) is another form of Cauchy's functional equation, and that g must have the form $g(s) = s^\alpha$, $\alpha \in \mathbb{R}$. Substituted into (15), we get

$$\tilde{\phi}_{\alpha,\tau}(s) = e^{-s^\alpha\tau}. \tag{20}$$

From the proof above we also get the form for the function $\psi(\gamma) = \gamma^\alpha$, for the dilation covariance property (9)

Corollary 1. *Let* $\{\phi_\tau | \tau \geq 0\}$ *be a convolution semigroup, where* ϕ_τ *are absolutely integrable causal functions, then the functions are dilation covariant iff*

$$D_\gamma \phi_\tau = \phi_{\gamma^\alpha \tau}. \tag{21}$$

We still need to determine for what values of the parameter α that $\tilde{\phi}_\tau$ is a Laplace transform of a normalizing non negative function. To be able to do this we need a few facts from Laplace transform theory.

Definition 2. *A function* f *on* $\mathbb{R}_+$ *is called* completely monotone *(see e.g. [2]), if it has derivatives of all orders and fulfill*

$$(-1)^n f^{(n)}(s) \geq 0, \quad s > 0. \tag{22}$$

Theorem 1 (Bernstein). *A function* f *on* $\mathbb{R}_+$ *is the Laplace transform of a non negative normalizing function, iff it is completely monotone and* $f(0) = 1$, *(see e.g. [2]).*

More specifically a function e^{-f} is completely monotone if f is a positive function with a completely monotone derivative.

From this we can prove our main theorem:

Theorem 2 (The form of causal scale space kernels). $\phi_{\alpha,\tau}(t)$ *is a causal scale space kernel iff*

$$\phi_{\alpha,\tau}(t) = \mathcal{L}^{-1}[e^{-s^\alpha \tau}] \tag{23}$$

for a fixed $0 < \alpha < 1$.

Proof. s^α is a positive function if $\alpha \geq 0$ and it is completely monotone if $0 \leq \alpha \leq 1$. $\phi_{0,\tau} = \delta(t)$ and $\phi_{1,\tau} = \delta(t - \tau)$ respectively, and thus continuity implies that $\alpha \neq 0$ and $\alpha \neq 1$. For $0 < \alpha < 1$ $\phi_{\alpha,\tau}(t)$ is continuous for $t \geq 0$.

There does not seem to be any known closed form of $\phi_{\alpha,\tau}$, but a series expansion is possible.

$$\phi_{\alpha,\tau}(t) = \mathcal{L}^{-1}[e^{-s^\alpha \tau}] \tag{24}$$

$$= \sum_{k=0}^{\infty} \mathcal{L}^{-1}[\frac{(-s^\alpha \tau)^k}{k!}] \tag{25}$$

$$= \frac{1}{t} \sum_{k=0}^{\infty} \frac{(-\tau)^k}{k! \Gamma(-k\alpha)} t^{-k\alpha}. \tag{26}$$

Remark 1. For the particular case $\alpha = 1/2$ an explicit form of the causal scale space kernel is known to be

$$\phi_{1/2,\tau}(t) = \frac{\tau}{\sqrt{4\pi}} \frac{\exp(-\tau^2/4t)}{t^{3/2}} = -2\partial_\tau k_t(\tau), \quad t \geq 0 \tag{27}$$

where k_σ is the Gaussian function

$$k_\sigma(x) = \frac{e^{-x^2}/4\sigma}{\sqrt{4\pi\sigma}} \tag{28}$$

Fig. 1. $\phi_{\alpha,1}(t)$ for $\alpha = \{0.3, 0.4, \mathbf{0.5}, 0.6, 0.7, 0.8, 0.9\}$, where larger α corresponds to functions peaking further to the right and the thicker line corresponds to $\alpha = 0.5$.

4 Stable Density Functions

The one parameter family of causal scale space kernels derived above, is well known in the field of probability theory as *extremal stable density functions* (the results from probability theory reviewed here can be found in e.g. [2]). These functions were one of the findings from the various attempts to generalize the central limit theorem during the first half of the 20:th century. The central limit theorem basically states that if $X_1, X_2, \ldots$ are mutually independent one dimensional random variables with zero mean and finite variance, the distributions of the normalized sums

$$S_n = F_n(X_1 + \ldots + X_n),$$

where F_n is a normalization function with appropriate properties, tend to a normal distribution as $n \to \infty$.

P. Levy generalized the central limit theorem by removing the requirements on zero mean and finite variance. For this generalization, the densities of the limit sums of stochastic variables are denoted stable density functions, which is a two parameter family of functions. One sided (causal) stable density functions forms a one parameter subfamily called extremal stable density functions, the same family of functions that we derived from the causal scale space kernel axioms in Definition 1. Symmetric stable density functions also form a one parameter subfamily of stable density functions, with the Cauchy and the Gaussian densities as notable members. The one parameter family of symmetric scale space kernel found by Pauwels et al. [8] is identical to the symmetric stable density functions.

Some of the relation between the generalized limit theorem and the causal scale space axioms (Definition 1) can roughly be sketched as follows: Density functions have unit area and are positive. Addition of stochastic variables correspond to convolution of their density functions. The dilation covariance requirement corresponds to properties of the normalization function in the limit theorems.

The stable density functions that are neither extremal nor symmetric can be said to lie between these two families and are more or less skew. If we were to define scale spaces (from the axioms in Definition 1 or the axioms in [8]), but without requiring causality or symmetry, the scale space kernels of those scale spaces would precisely correspond to the family of stable density functions.

4.1 Scaling Properties

All stable density functions except the Gaussian density function have infinite variance and the extremal stable density functions also have infinite mean. Stable density functions are also known to be unimodal. From causality and unimodality we can draw the conclusion that a temporal measurement gives the highest weight to values of the input signals that occurred a while ago, (which can also be seen in Figure 1. This can be described as that there is a certain delay for each measurements, and this delay will be larger for measurement at larger scales. The delay cannot be described in terms of the mean as the mean is infinite for causal scale space kernels. The mode i.e. the maximum of the density function can be used instead. For $\phi_{1/2,\tau}$ the mode $t_m(\tau)$ easily can be shown to be

$$t_m(\tau) = \frac{\tau^2}{6},$$

by solving $(d/dt)\phi_{1/2,\tau}(t) = 0$, for t. For general causal scale space kernels it seems harder to find the mode, but we can at least derive how the mode is a function of scale modulo a constant.

Proposition 1. *Let* $T_{\alpha,m} = t_{\alpha,m}(1)$ *then*

$$t_{\alpha,m}(\tau) = T_{\alpha,m}\tau^{1/\alpha}. \tag{29}$$

Proof. By setting $\gamma^\alpha \tau = 1$ that is $\gamma = \tau^{-1/\alpha}$ in Equation 21 we get

$$\phi_{\alpha,1}(T_{\alpha,m}) = \frac{1}{\gamma}\phi_{\alpha,\tau}(\frac{T_{\alpha,m}}{\gamma}) = \frac{1}{\gamma}\phi_{\alpha,\tau}(t_{\alpha,m}(\tau))$$

and therefore

$$t_{\alpha,m}(\tau) = \frac{T_{\alpha,m}}{\gamma} = T_{\alpha,m}\tau^{1/\alpha}.$$

5 Recursive Formulation

A theory about temporal measurement must respect causality, the observer cannot access the future. Furthermore, *the observer cannot access the past*. The only way the observer can use information from its past is through its memory of past measurements, the observer must embody its past. That is, some kind of memory must be involved in the model and an important property for a temporal measurement theory is its memory model. From the result above, where the temporal measurement is described in terms of a convolution of a kernel with the past signal, it might seem like the memory must contain the whole past signal. However, we will show that all information about the past that is needed is contained in the temporal scale space at the current moment. The scale space can be described in terms of an integrodifferential equation, where the temporal scale space evolves over time with the input signal as boundary condition.

5.1 Fractional Derivatives

The evolution equation involves fractional derivatives so we start by defining them.

Definition 3. *$D^p_{x,+}$, is called the p-order left sided Riemann-Liouville fractional derivative [12] in x and is defined as,*

$$D^p_{x,+}f(x) = \frac{1}{\Gamma(k-p)}\,(\partial_x)^k \int_0^x (x-y)^{k-p-1} f(y)\,dy, \quad (k-1 \le p < k) \tag{30}$$

$$= (\partial_x)^k \Big(\frac{x_+^{k-p-1}}{\Gamma(k-p)} * f(x)\Big), \tag{31}$$

where x_+ is equal to x for $x > 0$ and zero for $x \le 0$.

Fractional order derivatives are generalizations of ordinary derivatives. They are linear operators, satisfy a generalization of the Leibniz rule and integer order fractional derivatives are equivalent to ordinary derivatives.

Different families of fractional derivatives can be created by integrating over other intervals than in equation (30) [12]. We will also use left sided fractional derivatives in the sequel.

Definition 4. *$D^p_{x,-}$, is called the p-order right sided Riemann-Liouville fractional derivative [12] in x and is defined as,*

$$D^p_{x,-}f(x) = \frac{1}{\Gamma(k-p)}(-\partial_x)^k \int_x^\infty (y-x)^{k-p-1} f(y)\,dy, \qquad (k-1 \le p < k). \tag{32}$$

5.2 Infinitesimal Generator

Now we can state a partial integrodifferential equation that generates the causal scale spaces.

Theorem 3. *The temporal scale space $u(t,x) = (\phi_{\alpha,x} * f)(t), 0 < \alpha < 1$, is the unique solution to the partial integrodifferential equation*

$$\begin{cases} \partial_x u = -D^\alpha_{t,+} u \\ \lim_{x\to 0} u(t,x) = f(t) \end{cases} \tag{33}$$

Proof. A semigroup T, that fulfills a certain continuity requirement, *strong continuity*, is the unique solution to the *abstract Cauchy problem*, (see e.g. [4])

$$\begin{cases} \partial_t u = Au \\ \lim_{t\to 0} u(t) = f \end{cases} \tag{34}$$

A is denoted the *infinitesimal generator* of the semigroup, and is defined as $A = \lim_{h\to 0+} A_h$, where $A_h = (T(h) - I)/h$. A semigroup is strongly continuous if, for all f, $T(t)f$ is continuous in t on $\mathbb{R}_+$.

$\phi_{\alpha,\tau}$ is strongly continuous as a consequence of the continuity axiom (Definition 1). Its infinitesimal generator is readily found in the Laplace domain,

$$\mathcal{L}[A_h u] = \mathcal{L}[(\Phi_{\alpha,h} u - u)/h] = \frac{1}{h}(e^{-s^\alpha h} - 1)\mathcal{L}[u], \tag{35}$$

and

$$\mathcal{L}[A] = \lim_{h\to 0+} \frac{1}{h}(e^{-s^\alpha h} - 1) = -s^\alpha. \tag{36}$$

This is the Laplace transform of the left sided Riemann-Liouville fractional derivative of order α:

$$\begin{aligned}
\mathcal{L}[D^\alpha_{t,+}u] &= \mathcal{L}[(\partial_x)(\frac{x_+^{-\alpha}}{\Gamma(1-\alpha)} * u(x))] && (37)\\
&= s\mathcal{L}[\frac{x_+^{-\alpha}}{\Gamma(1-\alpha)} * u(x)] - \left(\frac{x_+^{-\alpha}}{\Gamma(1-\alpha)} * u(x)\right)_{x=0} && (38)\\
&= s\mathcal{L}[\frac{x_+^{-\alpha}}{\Gamma(1-\alpha)}]\mathcal{L}[u(x)] && (39)\\
&= ss^{-(1-\alpha)}\mathcal{L}[u(x)] && (40)\\
&= s^{-\alpha}\mathcal{L}[u(x)] && (41)\\
& && (42)
\end{aligned}$$

It should be noted that we still lack a time recursive formulation, as the fractional operator $D^\alpha_{t,+}$ for $0 < \alpha < 1$ applied on the temporal signal is non-local, (it has in fact support on the whole half axis).

5.3 Evolution Equation

Interestingly enough, the partial integrodifferential equation from Theorem 3 above can be transformed to an partial integrodifferential equation that only applies a first derivative on the temporal signal.

Theorem 4. *The temporal scale space $u(t,x) = (\phi_{\alpha,x} * f)(t), 0 < \alpha < 1$, is the unique solution to the partial integrodifferential equation*

$$\begin{cases} \partial_t u = D^{1/\alpha}_{x,-}u \\ \lim_{x\to 0} u(t,x) = f(t) \\ u(0,x) = 0. \end{cases} \tag{43}$$

Proof. We need a linear operator A_x in x that satisfies:

$$\partial_t u = A_x u. \tag{44}$$

In the Laplace transform domain this becomes:

$$\mathcal{L}[\partial_t u] = se^{-s^\alpha x}\mathcal{L}[f] = \mathcal{L}[A_x u] = A_x\mathcal{L}[u] = A_x[e^{-s^\alpha x}]\mathcal{L}[f], \tag{45}$$

and thus A_x must satisfy,

$$A_x e^{-s^\alpha x} = se^{-s^\alpha x}. \tag{46}$$

A linear operator with this property is the right sided Riemann-Liouville fractional derivative defined above, we have [12]:

$$D^\beta_{x,-}e^{-\lambda x} = \lambda^\beta e^{-\lambda x}, \qquad \Re\lambda > 0, \tag{47}$$

setting $\beta = 1/\alpha$ and $\lambda = s^\alpha$, we can see that $A_x = D^{1/\alpha}_{x,-}$ satisfies (46). It is also known that equations of the type (43) have a unique solution.

Theorem 4 shows that the temporal scale space can be described in terms of an evolution equation on the half line where the position corresponds to temporal scale and to older information, (larger scale gives higher weight to older information). The input signal is fed to $x = 0$ where the evolution equation only applies a local operation (the temporal first derivative) on the signal in the present moment and as a consequence we have found a time recursive formulation of the causal scale spaces. For this realization of causal scale spaces the temporal scale space is the only needed memory of earlier input, and the content of the memory diffuses over time.

For $\alpha = 1/2$ we have $D^{1/\alpha}_{x,-} = D^2_{x,-} = \partial^2_x$ and the evolution equation specializes to the *signaling equation*

$$\begin{cases} \partial_t u = \partial_x^2 u \\ \lim_{x\to 0} u(t,x) = f(t) \\ u(0,x) = 0. \end{cases} \tag{48}$$

The signaling equation describes how current is distributed in a semi infinite conductor when an temporally modulated electrical signal is applied at its end. It describes how heat is diffused in a semi infinite rod when a temporally modulated heat source is applied at its end, as well.

Theorem 5. *The signaling equation is unique among the evolution equations for causal scale spaces in the sense that it possesses both locality and positivity in the scale domain.*

Proof. As already noted, equation (33) always uses a non local operator in the temporal direction. The evolution equation (43), becomes a partial differential equation for when $1/\alpha$, $0 < \alpha < 1$ is an integer i.e. for $\alpha = 1/k$, $k \geq 2$, where k is an integer. The locally generated evolution equations therefore has the form $\partial_t u = \partial_x^k u$, $k \geq 2$. And this equation has only a positive Greens function for $k = 2$, i.e. for $\alpha = 1/2$.

6 Discretization of Causal Scale Spaces

Both the partial integrodifferential equations (33) and (43) can be numerically implemented by using discretization of fractional derivatives from e.g. [9]. Discretizations of fractional derivatives need to be computed for quite a large number of grid points to achieve a reasonable low numerical error. The locally generated signaling equation (48) can be much more efficiently implemented and we will focus on finding a numerical implementation for it.

An important step in finding a numerical implementation of the temporal scale space is to find a suitable discretization of the problem. Florack [3] states that the natural discretization of a space, is such that the grid steps are constant in the natural parametrization of the Lie group that generates the space. As we have decided that time is translation covariant we obtain constant intervals in the temporal direction. The scale is considered to be scale covariant which leads to a geometric progression of grid points in the scale direction.

6.1 Discrete Second Derivative on Log Spaced Grid

We need to derive a discretization of the second derivative operation for a grid with geometrical progression to be able to compute the signaling equation on such a grid.

We denote grid points by x_i and use the following notation.

$$u_i = u(x_i)$$
$$\varDelta_{i+1} = x_{i+1} - x_i$$

Theorem 6 (Saulyev [13]). *For general non uniform grids the second derivative becomes:*

$$\frac{\partial^2 u_i}{\partial x^2} = \frac{2u_{i+1}}{\varDelta_{i+1}(\varDelta_{i+1} + \varDelta_i)} + \frac{2u_{i-1}}{\varDelta_i(\varDelta_{i+1} + \varDelta_i)} - \frac{2u_i}{\varDelta_i \varDelta_{i+1}} + o(\varDelta_i), \tag{49}$$

where the error term in general is of linear order in $\varDelta_i$.

If we specialize the above theorem for a grid with geometric progression:

$$x_i = x_0 h^i$$

where $h > 1$ we get an error term of order $o((h-1)^2)$.

Theorem 7. *For grids with geometric progression the second derivative becomes:*

$$\frac{\partial^2 u_i}{\partial x^2} = \frac{2}{x_i^2 (h-1)(h-\frac{1}{h})}(u_{i+1} - (h+1)u_i + hu_{i-1}) + o((h-1)^2), \tag{50}$$

where the error term is of quadratic order in $(h-1)$, if $(h-1)$ is small enough.

Proof. For a grid with geometric progression we have that:

$$\varDelta_{i+1} = x_i h - x_i = x_i(h-1)$$
$$\varDelta_i = x_i - x_i/h = x_i(1 - \frac{1}{h}).$$

Inserting this in equation (49), we get:

$$\begin{aligned}
\delta_i^2 u &= \frac{2u_{i+1}}{\varDelta_{i+1}(\varDelta_{i+1} + \varDelta_i)} + \frac{2u_{i-1}}{\varDelta_i(\varDelta_{i+1} + \varDelta_i)} - \frac{2u_i}{\varDelta_i \varDelta_{i+1}} \\
&= \frac{2}{\varDelta_{i+1}(\varDelta_{i+1} + \varDelta_i)} \left(u_{i+1} + \frac{\varDelta_{i+1}}{\varDelta_i} u_{i-1} - \frac{\varDelta_{i+1} + \varDelta_i}{\varDelta_i} u_i \right) \\
&= \frac{2}{x_i^2 (h-1)(h-1/h)} \left(u_{i+1} + \frac{h-1}{1-1/h} u_{i-1} - \frac{h-1/h}{1-1/h} u_i \right) \\
&= \frac{2}{x_i^2 (h-1)(h-1/h)} \left(u_{i+1} + hu_{i-1} - (h+1)u_i \right).
\end{aligned}$$

For the error term we have:

$$\begin{aligned}
e(h-1) &= \sum_{j=3}^{\infty} \frac{2}{j!} \frac{\Delta_{i+1}^{j-1} + (-1)^j \Delta_i^{j-1}}{\Delta_{i+1} + \Delta_i} \frac{\partial^j u_i}{\partial x^j} \\
&= \sum_{j=3}^{\infty} \frac{2}{j!} \frac{(h-1)^{j-1} + (-1)^j (1-1/h)^{j-1}}{(h-1) + (1-1/h)} x_i^{j-2} \frac{\partial^j u_i}{\partial x^j} \\
&= \sum_{j=3}^{\infty} \frac{2(h-1)^{j-2}}{j!} \frac{h^{j-1} + (-1)^j}{h+1} x_0 \frac{\partial^j u_i}{\partial x^j} \\
&= \sum_{j=3,5,7,\ldots} \frac{2(h-1)^{j-1}}{j!} \frac{\sum_{k=0}^{j-2} h^k}{h+1} x_0 \frac{\partial^j u_i}{\partial x^j} + \\
&= \sum_{j=4,6,8,\ldots} \frac{2(h-1)^{j-2}}{j!} \frac{h^{j-1} + 1}{h+1} x_0 \frac{\partial^j u_i}{\partial x^j} \\
&= o((h-1)^2).
\end{aligned}$$

7 Numerical Scheme for the Signaling Equation

Now we have what is needed for formulating a numerical scheme for the signaling equation. First we recall that for an explicit solution of the diffusion equation, $\Delta_t/\Delta_x^2 < 1/2$ must hold for the solution to be numerically stable (see e.g. [10]). This is fairly unattractive for the causal scale space as it means that for a given temporal sampling of the input signal there is a lower limit for how fine sampling we can choose in the scale domain. It therefore seems to be better to use an implicit numerical solution as it always is stable (see e.g. [10] for details about implicit solutions of the heat equation) .

For the temporal derivation the discretization

$$\delta_t u(t,x) = u(t,x) + \frac{3}{2} u(t-\Delta_t, x) - \frac{1}{2} u(t-2\Delta_t, x),$$

is a good choice as it has quadratic stability $e = o(\Delta_t^2)$ and only uses past grid points. The proposed numerical scheme has second order stability both in time and scale and can be implemented with four additions, four multiplications and two divisions per grid point.

Fig. 2. Numerical experiment, the signal is in the foreground and the logarithm of the scale increase away from the viewer. The time is directed to the right.

8 Discussion

If we require the measurement kernels to both respect temporal causality and scale covariance, the maximum (or the mean if it exists) will move backwards in time with increasing scale. Therefore we will never be able to measure what happens at the current moment, a measurement on a fine scale will reflect an event that happened just a short while ago while a measurement on a coarser scale will describe something that happened further back in time. A temporal measurement thus involves *two* different points in time: the one the measurement is performed at, t_0 and the one that has largest influence on the measurement t_m. The distance between these point is a function of scale $t_0 - t_m = f(\tau)$, the *influence curve*. From scale covariance considerations f typically should be on the form $f(\tau) = \beta\tau^\alpha$ for some $\alpha, \beta > 0$. Compare this with the situation for spatial measurements: reflectional or rotational symmetry means that the point where the measurement is performed also is the point that has maximal influence on the measurement.

Some earlier axiomatization of temporal scale spaces [5, 7] have required non-creation of structure along the scale dimension. In Koenderink's axiomatization there were no solutions fulfilling temporal causality. He solved this by doing a remapping of the temporal dimension before applying ordinary scale space. In Lindeberg's and Fagerström's axiomatization there were solutions but they lacked scale covariance and had to be discrete in either time or space. From the above considerations about how larger scale leads to the measurement of events earlier in time the requirement of non-creation of structure along the scale dimension seems to be an unnecessarily strong. It might be more fruitful to require non-creation of structure along the influence curve instead.

Koenderink motivates the logarithmic re-mapping of the time axis by analogy to our memories: we have finer temporal resolution on events taking place seconds ago than on events years ago. As already indicated, if we want the temporal scale space theory to be about measurement, we have to make careful distinction between the actual measurements and the memory of them. While a logarithmic mapping of time in the *memory* domain seem to be good first approximation of an uncommitted memory, we believe that the actual measurement process should be the same at every instant of time, i.e. it should be translationally invariant in time.

References

1. Aczél, J., Dhombres., J.: Functional Equations in Several Variables. Encyclopaedia of Mathematics and its Applications. Cambridge University Press (1989)
2. Feller, W.: An introduction to probability theory and its application, volume 2. John Willey & Sons, Inc. (1966)
3. Florack, L.M.J.: Image Structure. Series in Mathematical Imaging and Vision. Kluwer Academic Publishers, Dordrecht, Netherlands (1997)
4. Hille, E., Phillips, R.S.: Functional analysis and semi-groups. American Mathematical Society (1957)

5. Koenderink, J.J.: Scale-time. Biological Cybernetics **58** (1988) 169–162
6. Koenderink, J.J., Kappers, A., van Doorn, A.J.: Local operations: The embodiment of geometry. Orban, G.A., Nagel, H.H. (Eds), Artificial and Biological Vision Systems, Basic Research Series, Springer Verlag (1992) 1–23
7. Lindeberg, T., Fagerstrm, D.: Scale-space with causal time direction. Proc. 4th European Conference on Computer Vision, volume **1064**, Cambridge, UK, Springer Verlag, Berlin (1996) 229–240
8. Pauwels, E.J., VanGool, L.J., Fiddelaers, P., Moons, T.: An extended class of scale-invariant and recursive scale space filters. PAMI, **17**(7) (1995) 691–701
9. Podlubny, I.: Fractional Differential Equations. Academic Press (1999)
10. Richtmyer, R.D., Morton, K.W.: Difference Methods for Initial-Value Problems. Interscience Publishers, 2 ed. (1967)
11. Salden, A.H., Haar Romeny, B.M. ter, Viergever, M.A.: Linear scale-space theory from physical principles. JMIV, **9**(2) (1998) 103–139
12. Samko, S.G., Kilbas, A.A., Marichev, O.I.: Fractional integrals and derivatives : theory and applications. Gordon and Breach Science Publishers, cop., Yverdon (1992)
13. Saulyev, V.K.: Integration of Equatioons of Parabolic Type by the Method of Nets. Pergamon Press (1964)
14. Treves, F.:. Topological Vector Spaces, Distributions and Kernels. Academic Press (1967)
15. Weickert, J., Ishikawa, S., Imiya, A.: On the history of gaussian scale-space axiomatics. Gaussian Scale-Space Theory (1997) 45–59