universität innsbruck

Lecture 6
Motion II: Spatiotemporal filters

703142. Computer Vision

Assoz.Prof. Antonio Rodríguez-Sánchez, PhD.

Outline

- Introduction
- Intuition
- Orientation in visual space-time
- A representation for spatiotemporal patterns
- Spatiotemporal boundaries
- Application to detection and tracking

Introduction

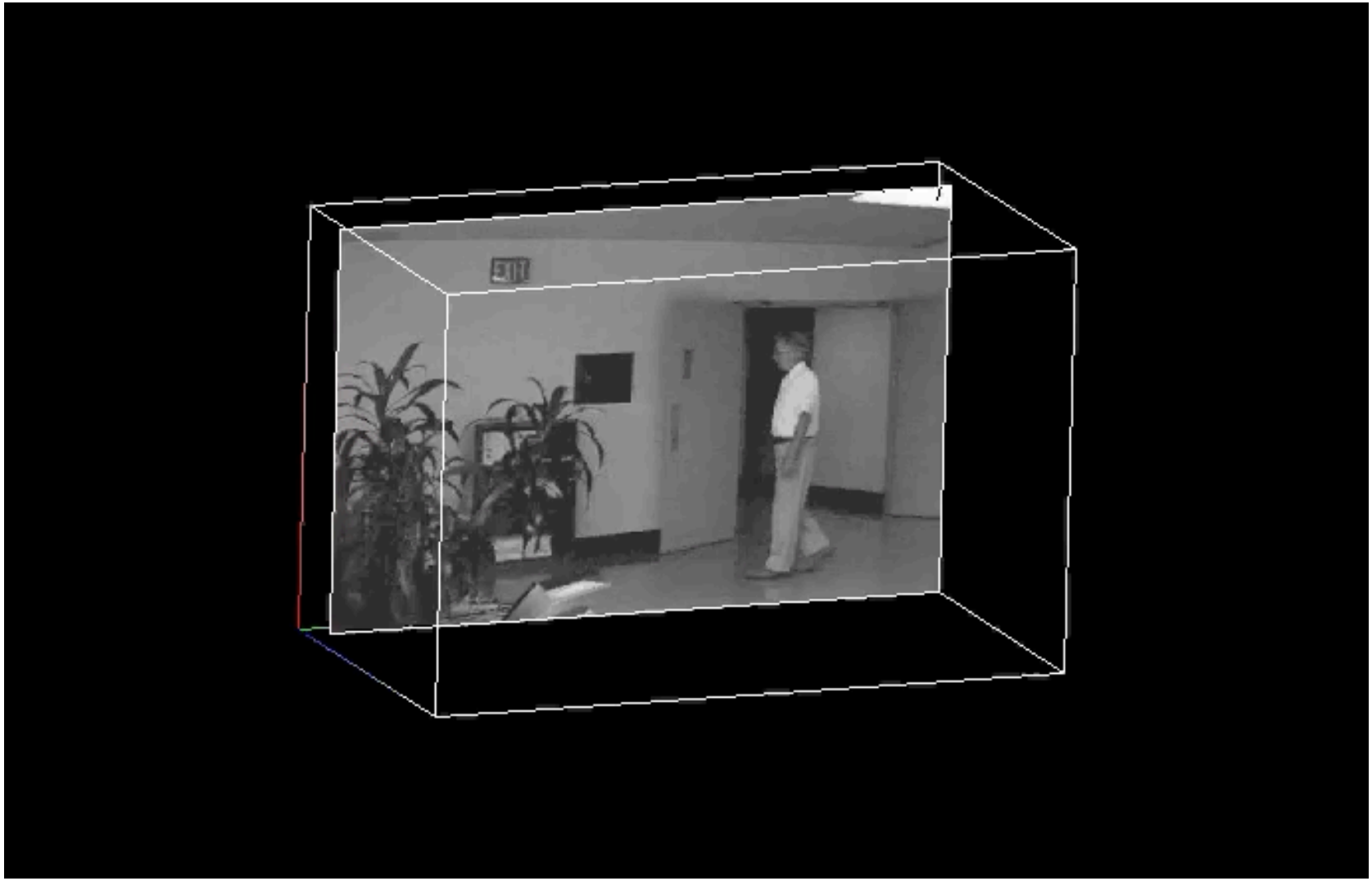
- We have considered the analysis of spatial structure.
  - Oriented, bandpass representations.
- We have considered analysis of the temporal dimension
  - Motion
- Now we consider the integrated analysis and interpretation of the spatial and temporal dimensions.
  - Spatiotemporal analysis

Outline

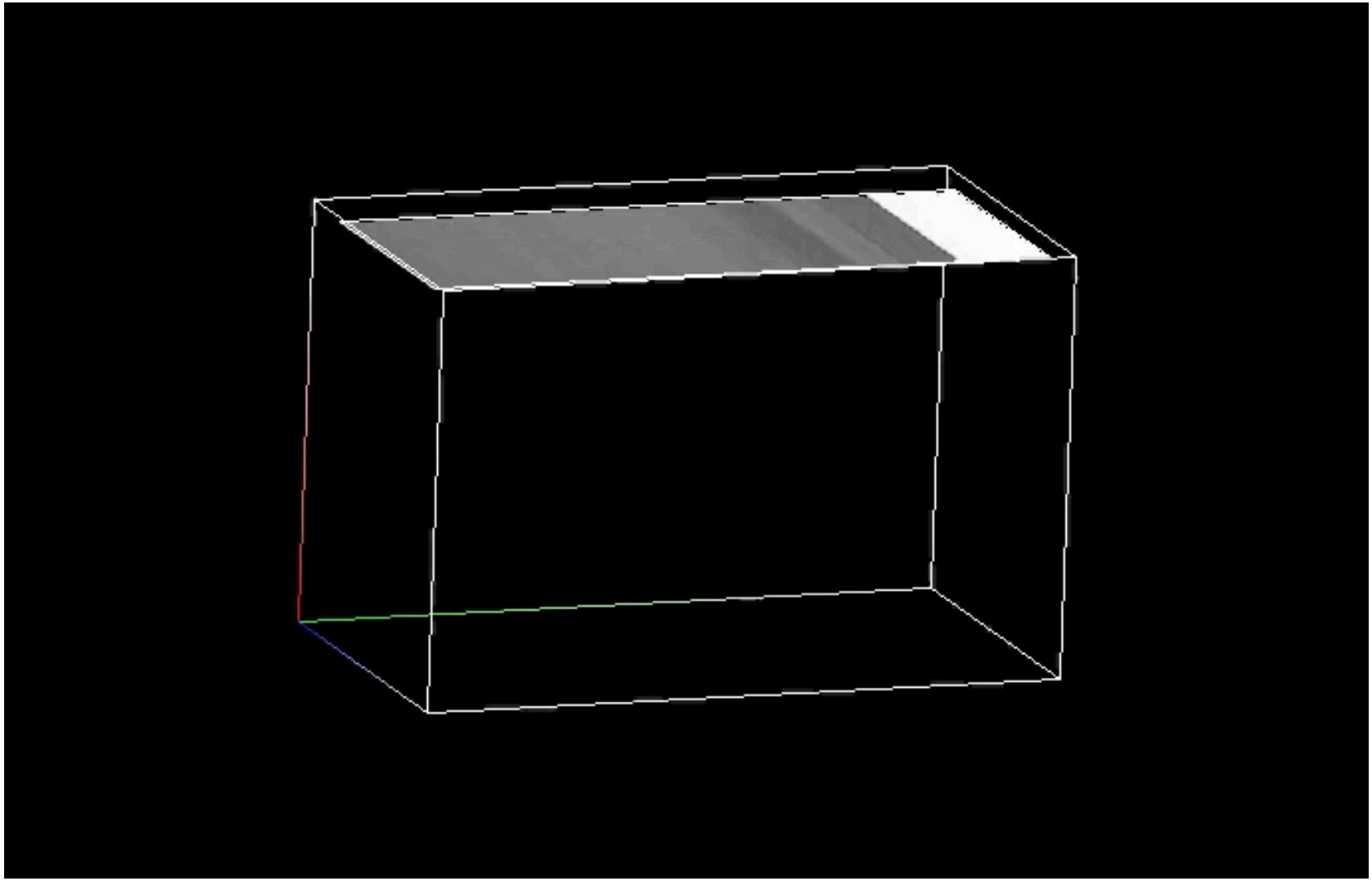
- Introduction
- Intuition
- Orientation in visual space-time
- A representation for spatiotemporal patterns
- Spatiotemporal boundaries
- Application to detection and tracking

Intuition

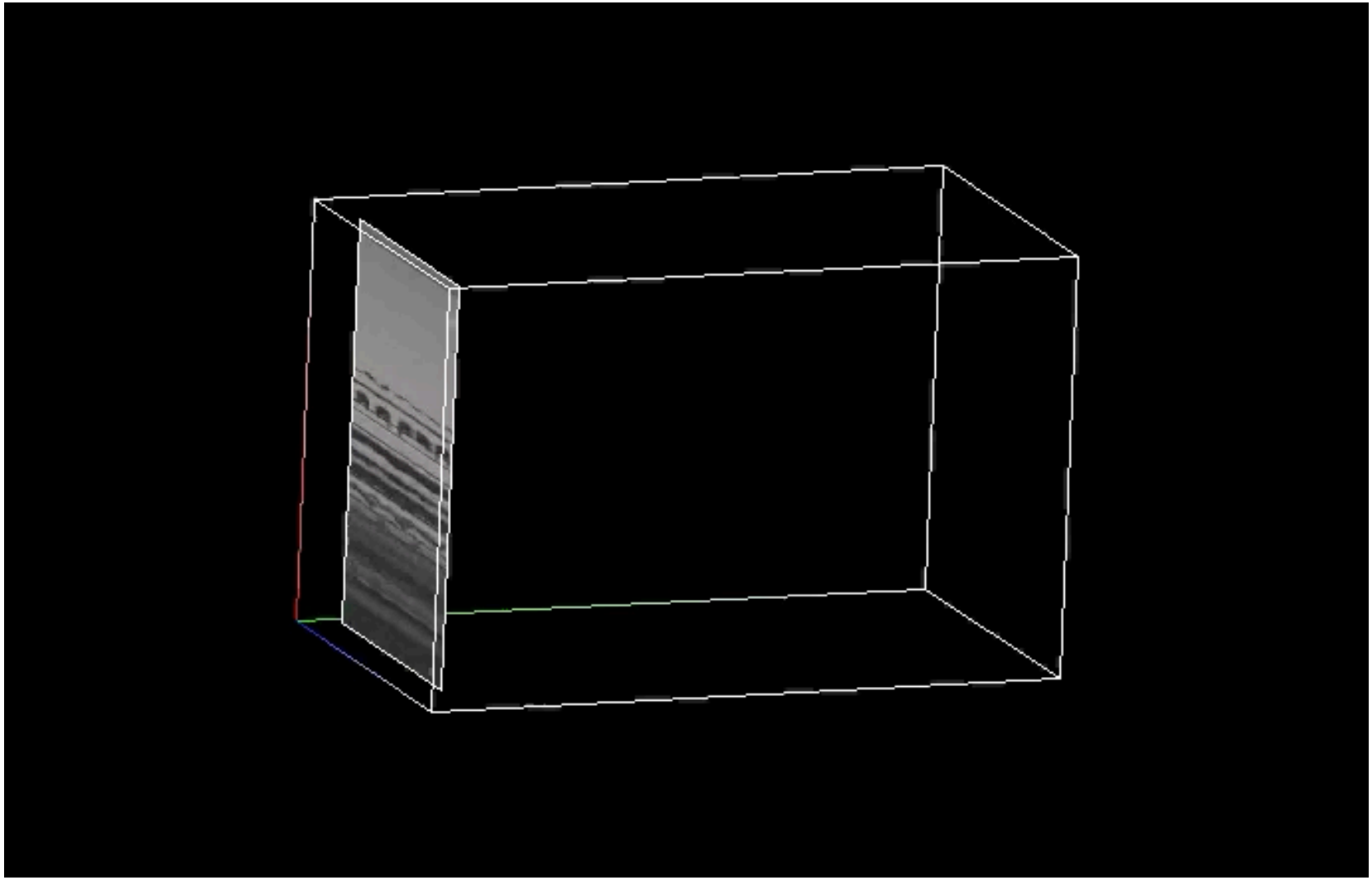
Intuition: XY



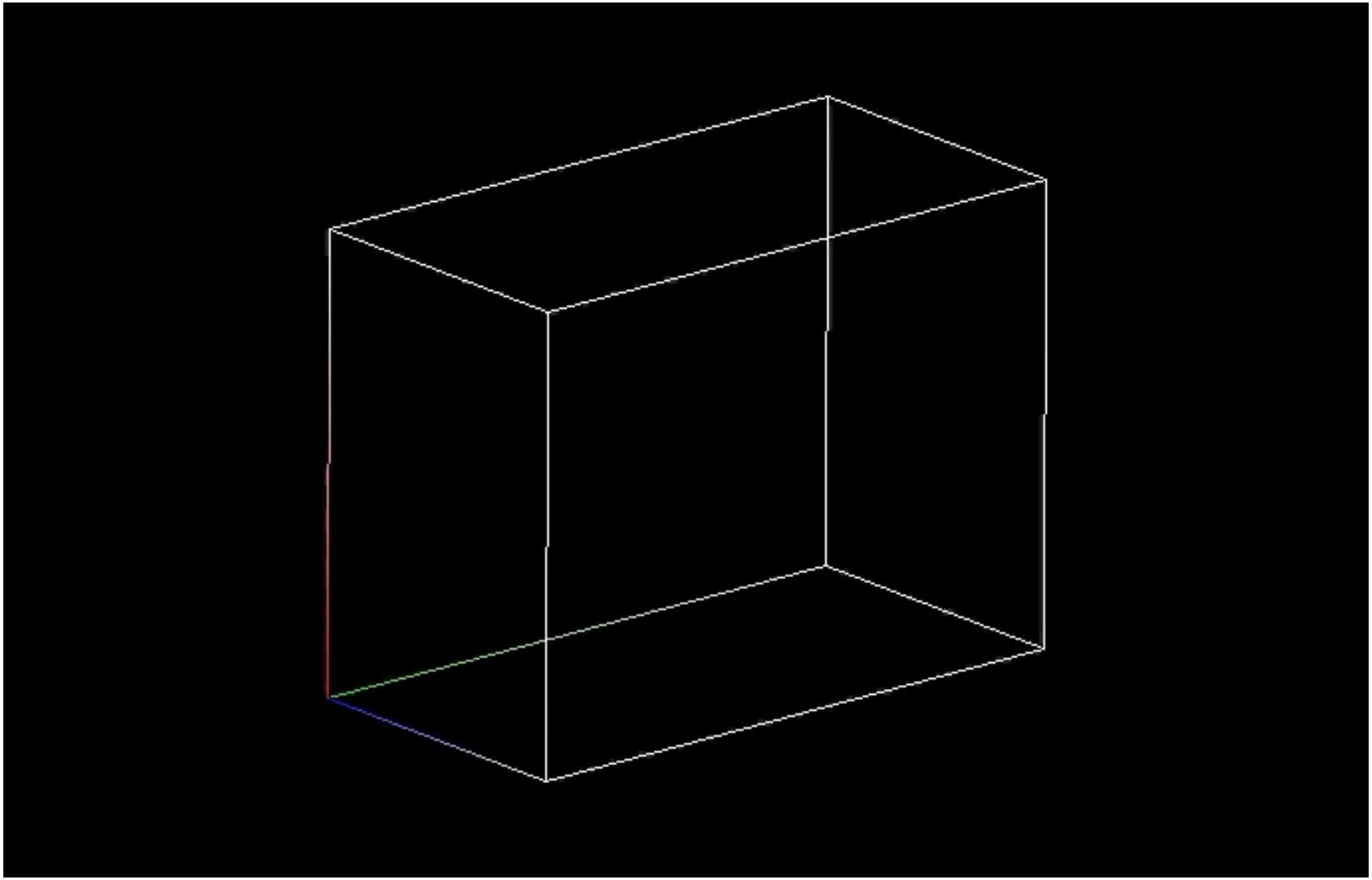
Intuition: XT



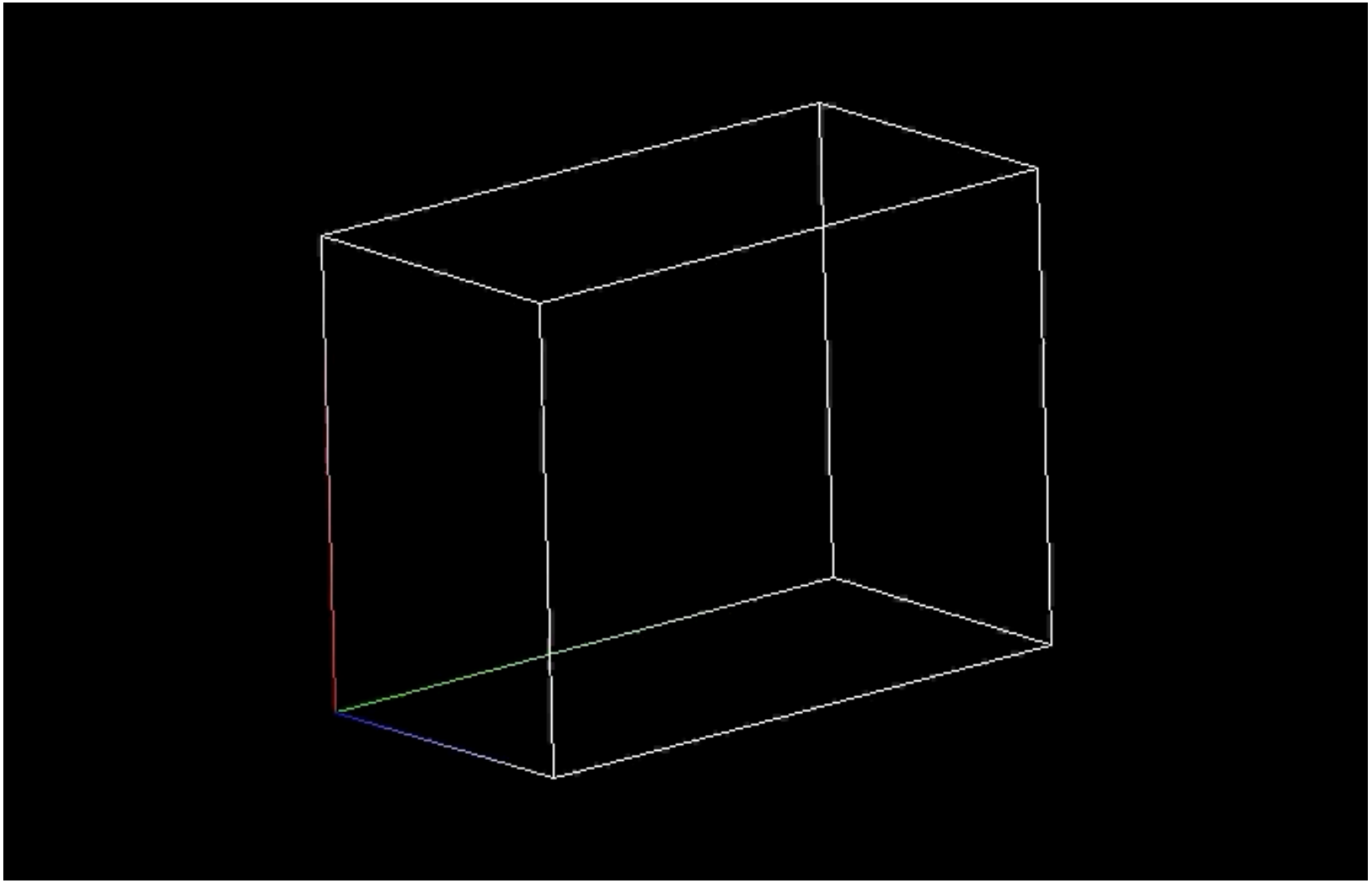
Intuition: YT



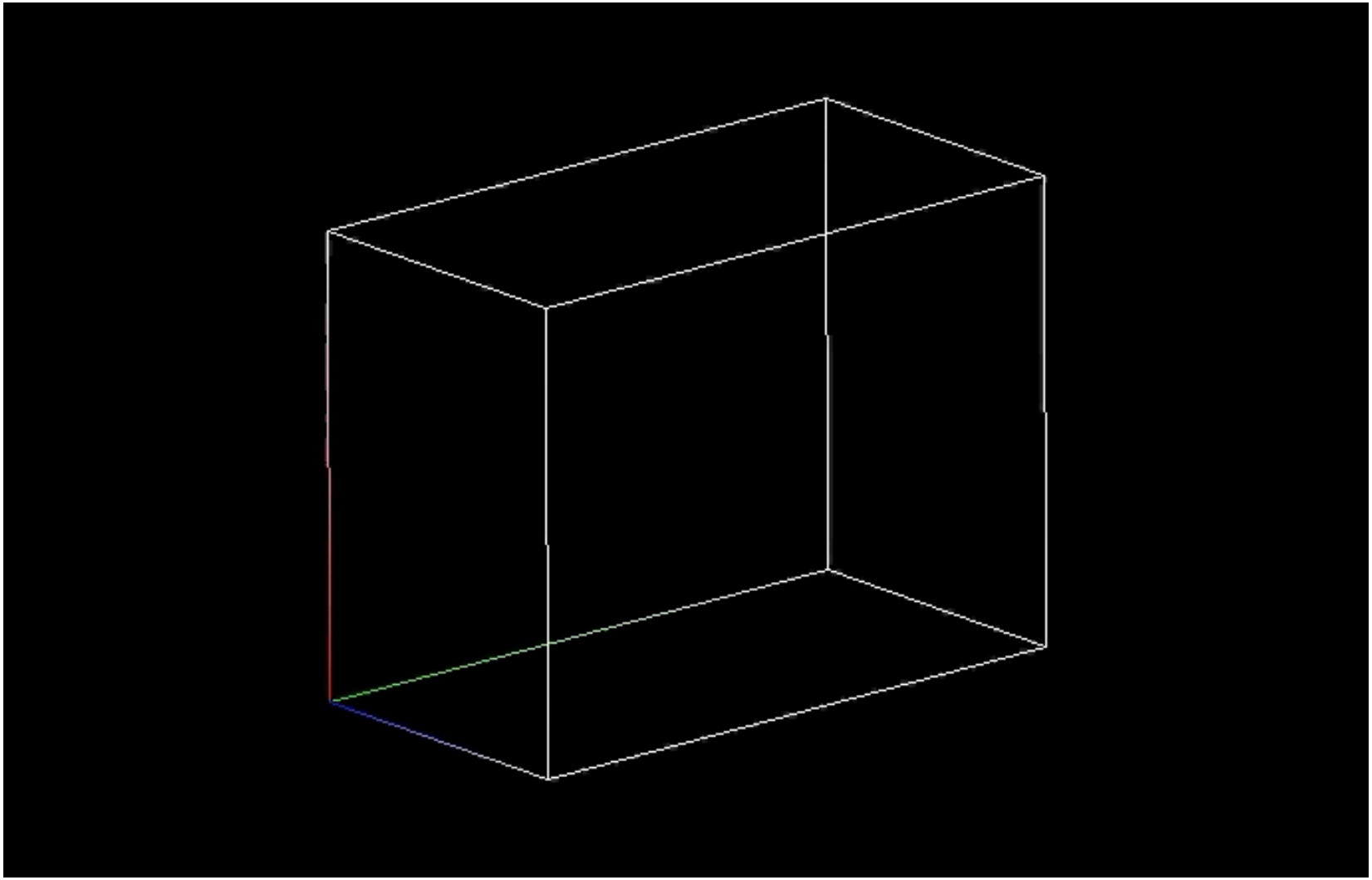
Intuition: Another example



Intuition: Another example



Intuition: Another example



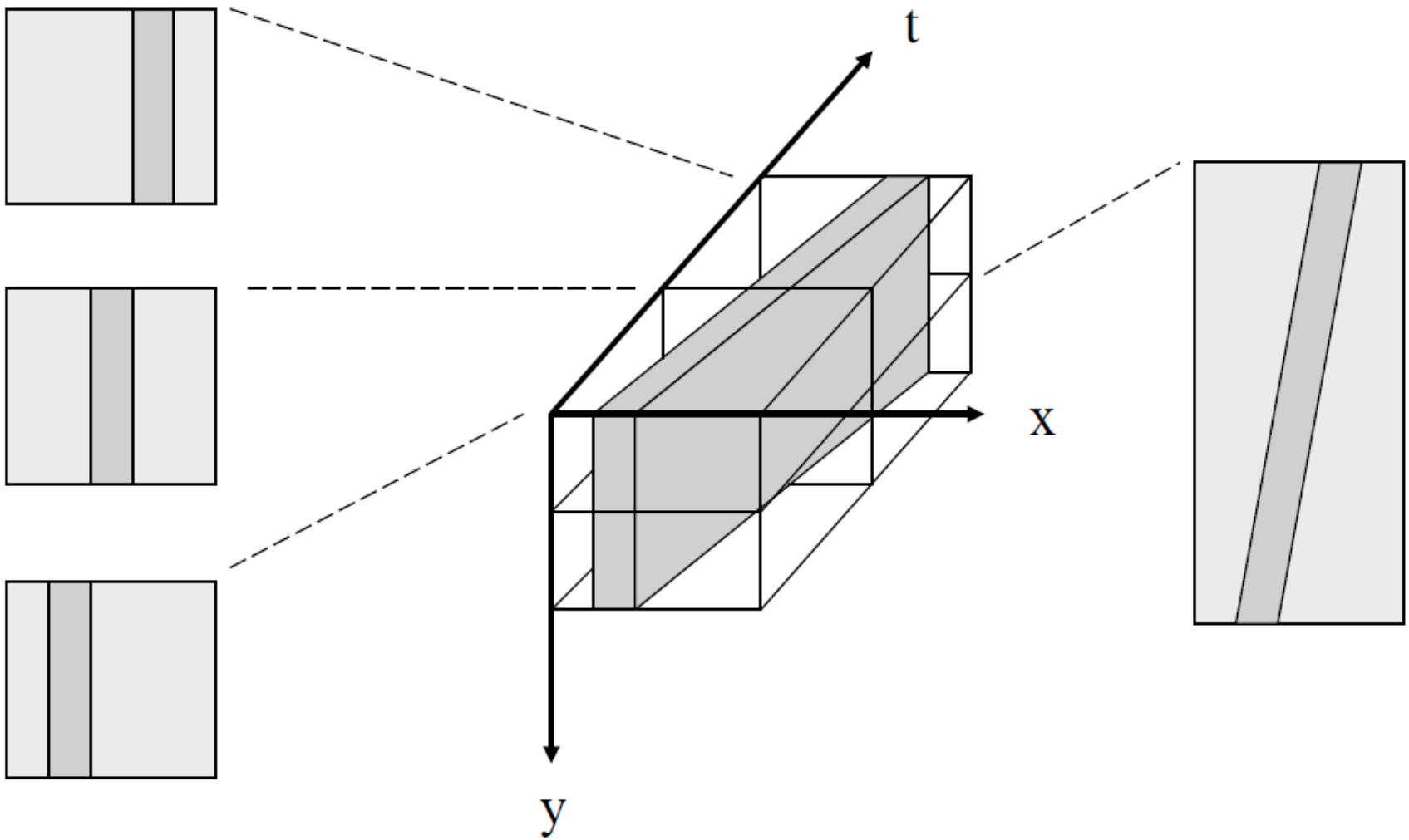
Outline

- Introduction
- Intuition
- Orientation in visual space-time
- A representation for spatiotemporal patterns
- Spatiotemporal boundaries
- Application to detection and tracking

Orientation in visual space-time

- The local orientation (or lack thereof) of a pattern is one of its most salient characteristics.
- Geometrically, orientation captures the local first-order structure of a pattern.
- For vision, local spatiotemporal orientation can have additional interpretations.
  - Image velocity is manifest as spatiotemporal orientation.
  - And more...

Orientation in visual space-time

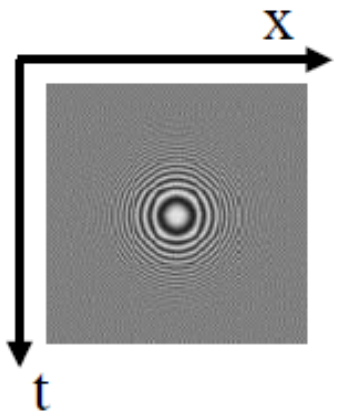


Orientation in visual space-time

- Goal is to analyze spatiotemporal data according to its local orientation structure.
- Consider orientation in x-t and y-t planes, with local weighted averaging in orthogonal spatial dimension.
- Choose a representation with multiple bands each tuned for certain orientations in a spatiotemporal plane, R, L, S, F and U, D, S, F.
- Consider single spatiotemporal scale (for now).

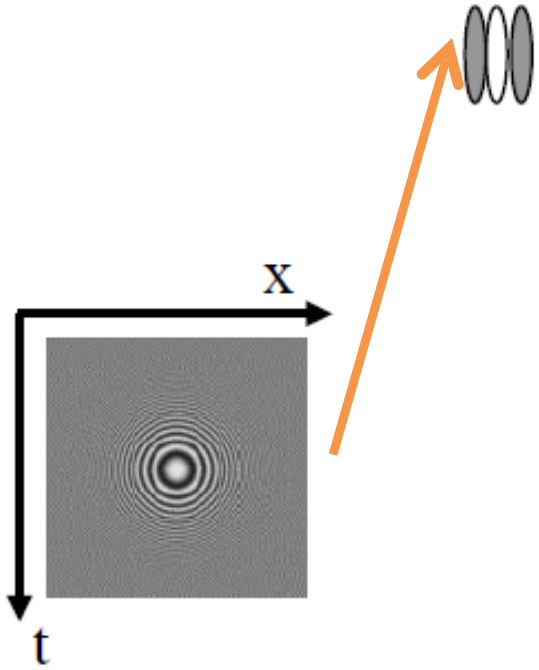
Orientation in visual space-time

- Apply filters tuned to 4 different orientations in both *x-t* and *y-t* domains.
  - In general, might consider additional directions.
- Filter specifics:
  - Oriented bandpass filters in spatiotemporal slice.
  - Lowpass filter in orthogonal spatial dimension.
  - Pointwise squared to yield local “oriented energy”.



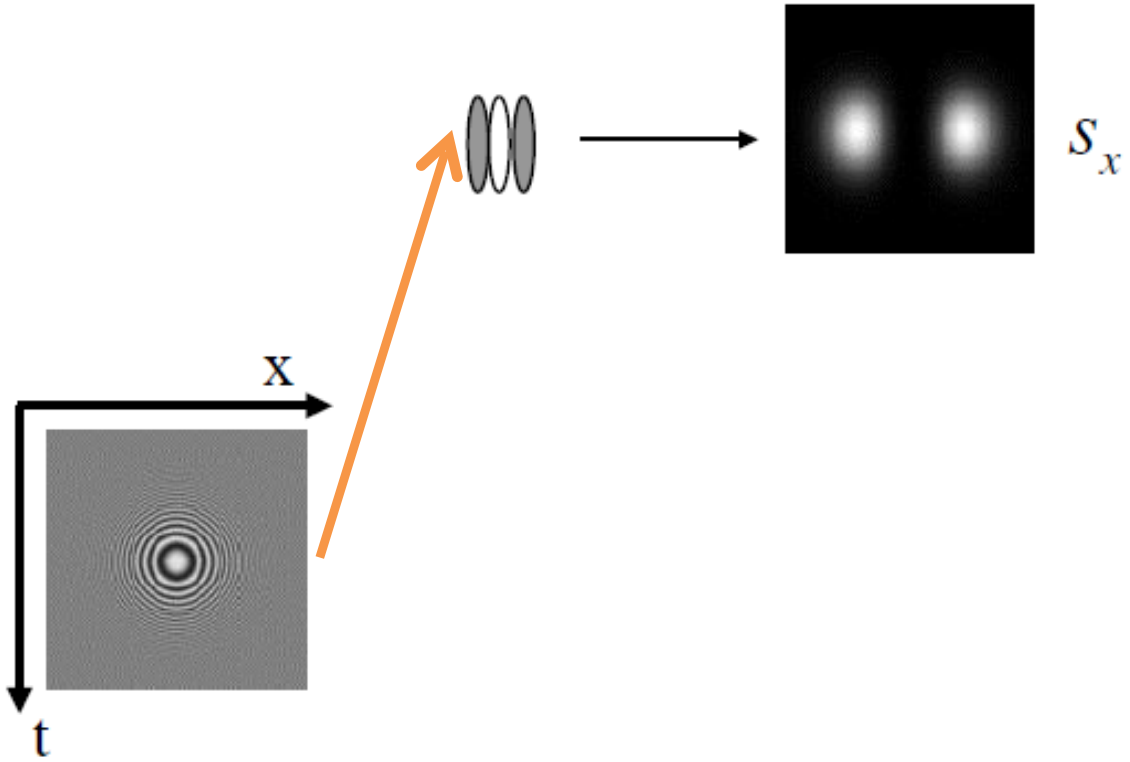
Orientation in visual space-time

- Apply filters tuned to 4 different orientations in both *x-t* and *y-t* domains.
  - In general, might consider additional directions.
- Filter specifics:
  - Oriented bandpass filters in spatiotemporal slice.
  - Lowpass filter in orthogonal spatial dimension.
  - Pointwise squared to yield local “oriented energy”.



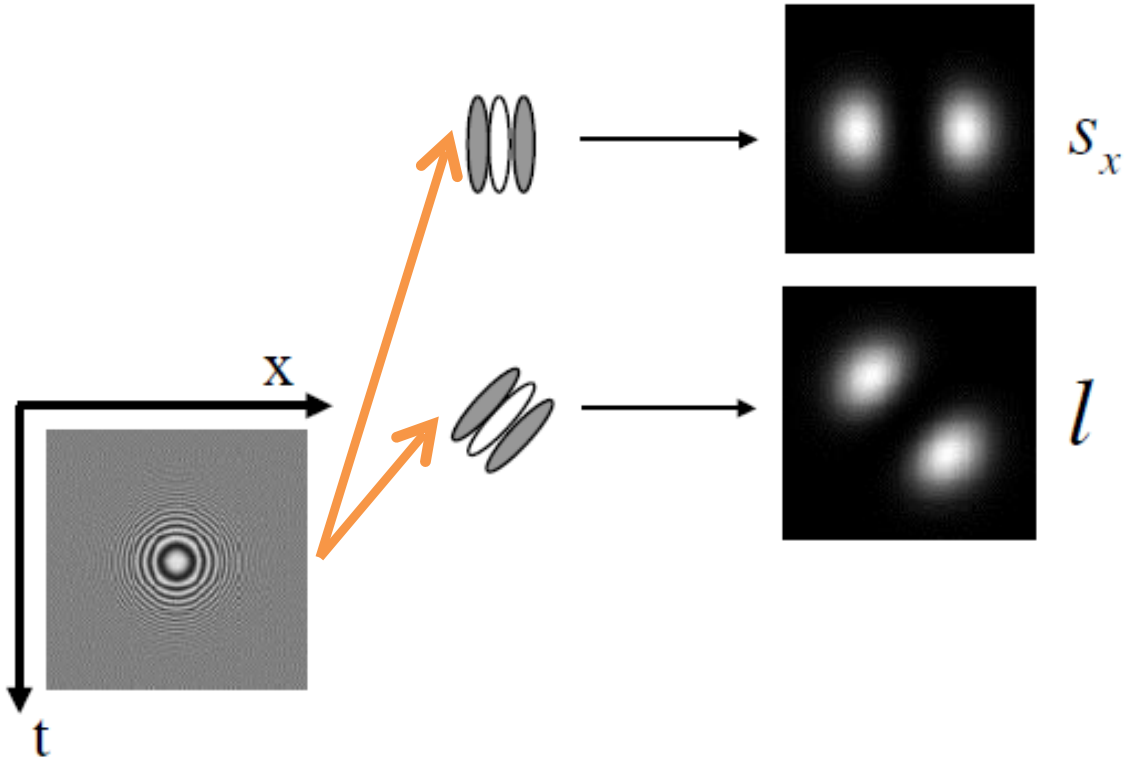
Orientation in visual space-time

- Apply filters tuned to 4 different orientations in both *x-t* and *y-t* domains.
  - In general, might consider additional directions.
- Filter specifics:
  - Oriented bandpass filters in spatiotemporal slice.
  - Lowpass filter in orthogonal spatial dimension.
  - Pointwise squared to yield local “oriented energy”.



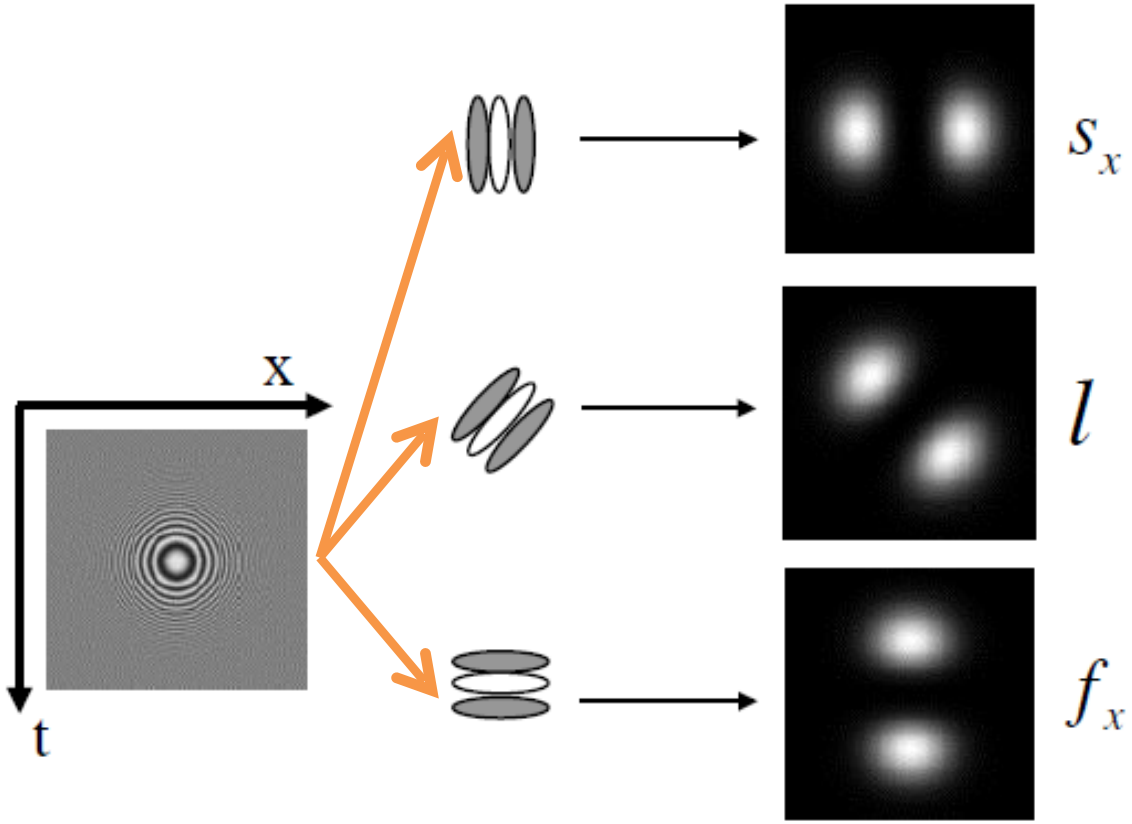
Orientation in visual space-time

- Apply filters tuned to 4 different orientations in both *x-t* and *y-t* domains.
  - In general, might consider additional directions.
- Filter specifics:
  - Oriented bandpass filters in spatiotemporal slice.
  - Lowpass filter in orthogonal spatial dimension.
  - Pointwise squared to yield local “oriented energy”.



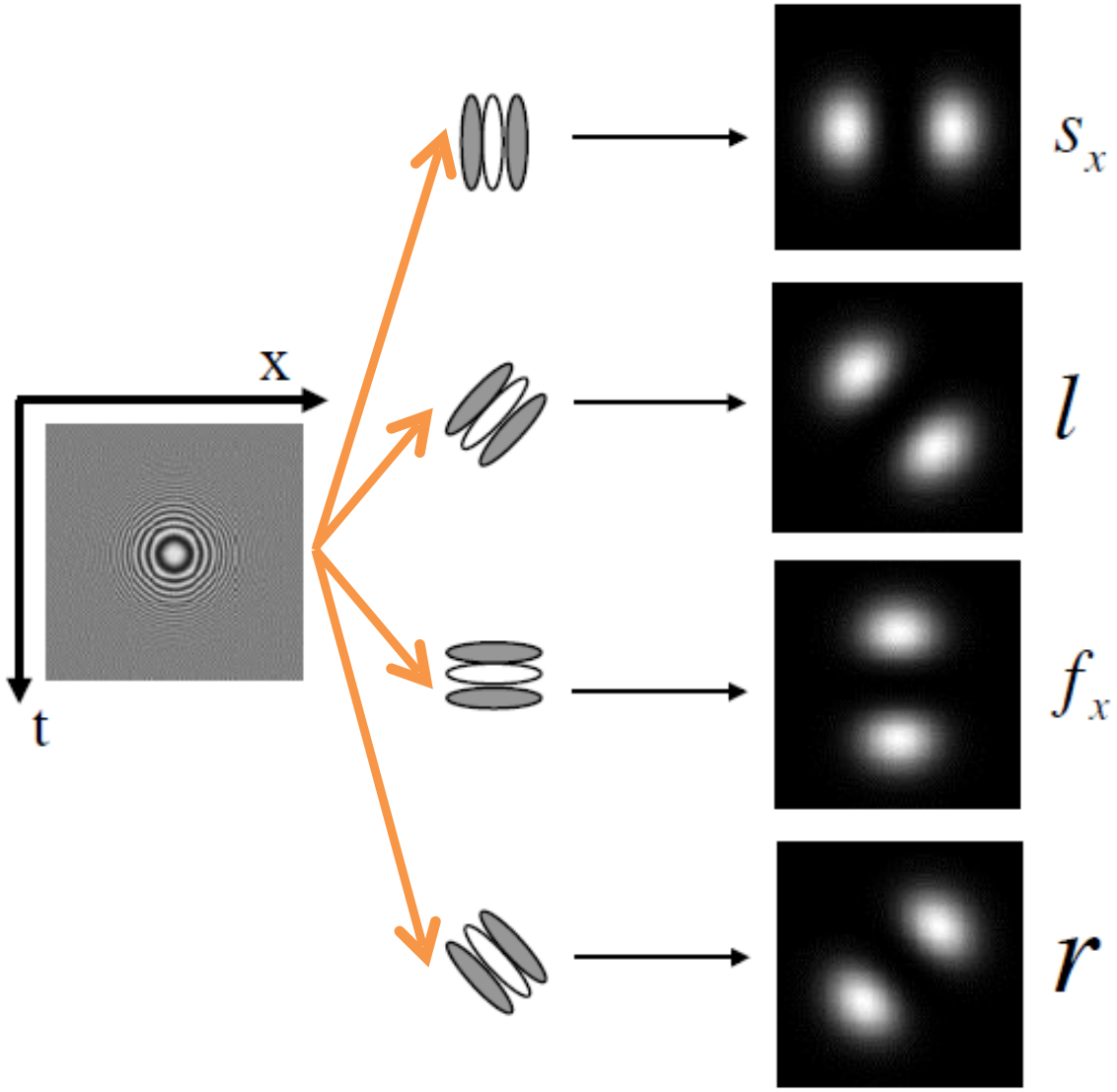
Orientation in visual space-time

- Apply filters tuned to 4 different orientations in both *x-t* and *y-t* domains.
  - In general, might consider additional directions.
- Filter specifics:
  - Oriented bandpass filters in spatiotemporal slice.
  - Lowpass filter in orthogonal spatial dimension.
  - Pointwise squared to yield local “oriented energy”.



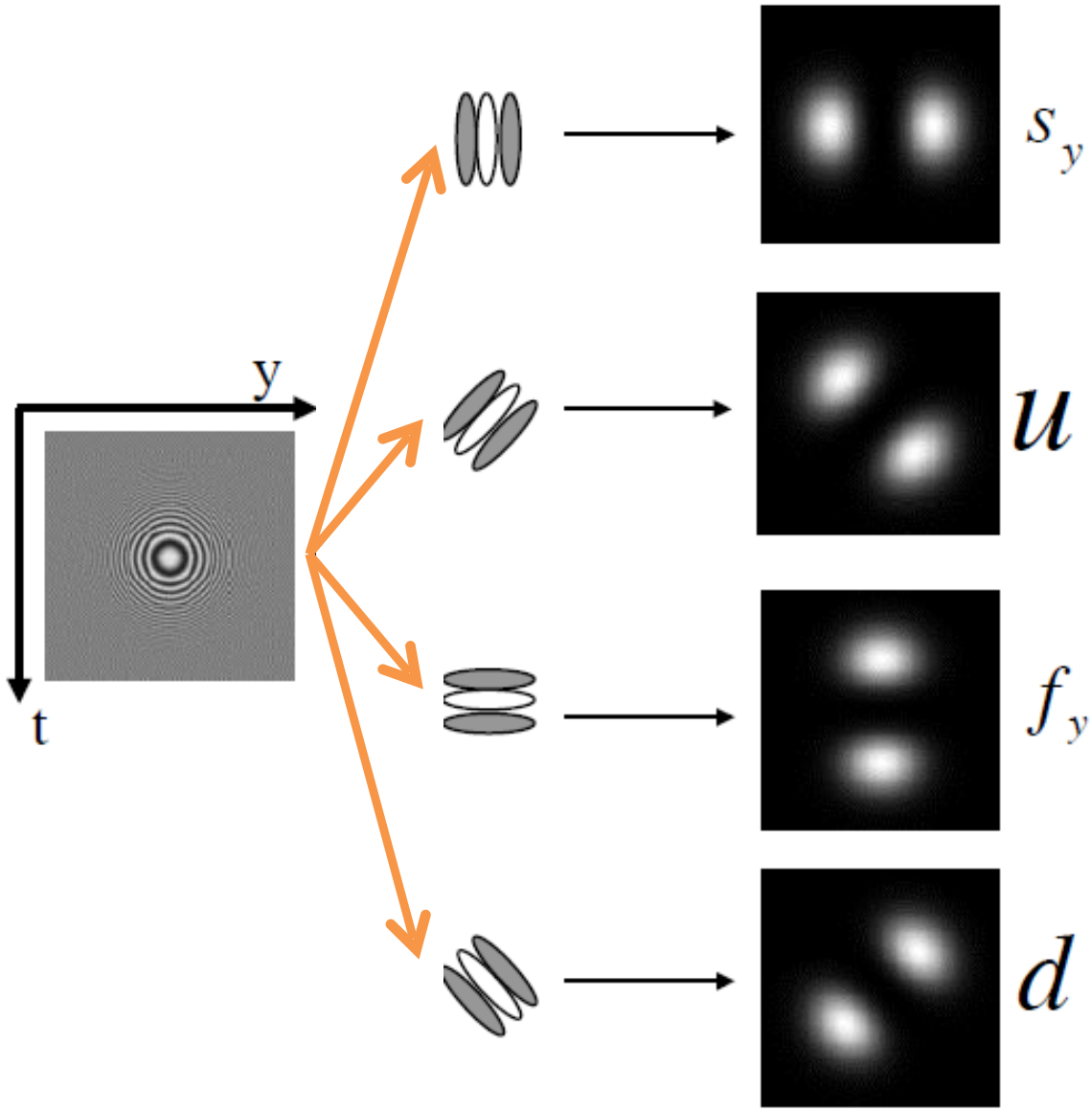
Orientation in visual space-time

- Apply filters tuned to 4 different orientations in both *x-t* and *y-t* domains.
  - In general, might consider additional directions.
- Filter specifics:
  - Oriented bandpass filters in spatiotemporal slice.
  - Lowpass filter in orthogonal spatial dimension.
  - Pointwise squared to yield local “oriented energy”.



Orientation in visual space-time

- Apply filters tuned to 4 different orientations in both *x-t* and *y-t* domains.
  - In general, might consider additional directions.
- Filter specifics:
  - Oriented bandpass filters in spatiotemporal slice.
  - Lowpass filter in orthogonal spatial dimension.
  - Pointwise squared to yield local “oriented energy”.



Orientation in visual space-time

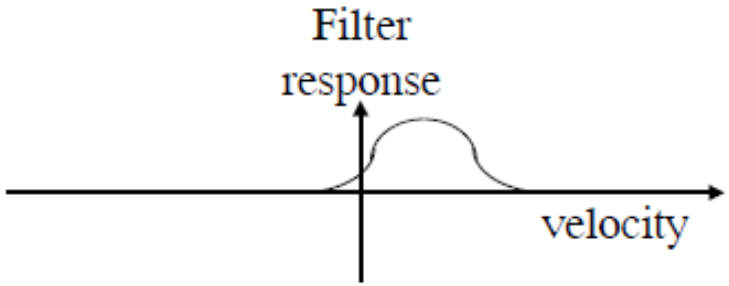
- Normalization
  - For any given orientation, the filter response is a joint function of orientation and contrast.
- Normalization yields purer measure of orientation

$$R(x,y,t) = \frac{r(x,y,t)}{r(x,y,t) + l(x,y,t) + s_x(x,y,t) + f_x(x,y,t) + \varepsilon}$$

- With ε a small bias added for stability
- Similarly for l, s_x, f_x and their y-t counterparts

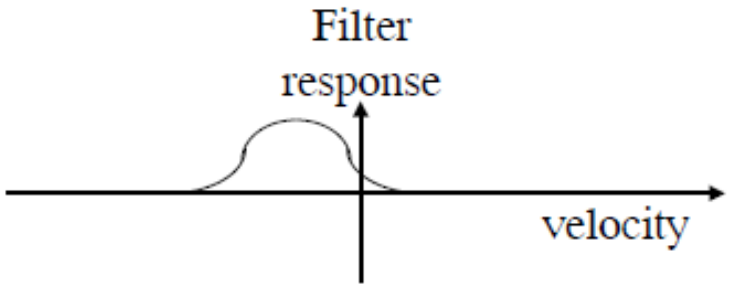
Orientation in visual space-time

- Velocity
  - Consider the response to a horizontally moving pattern of ***r***, ***l*** and ***s*** filters.



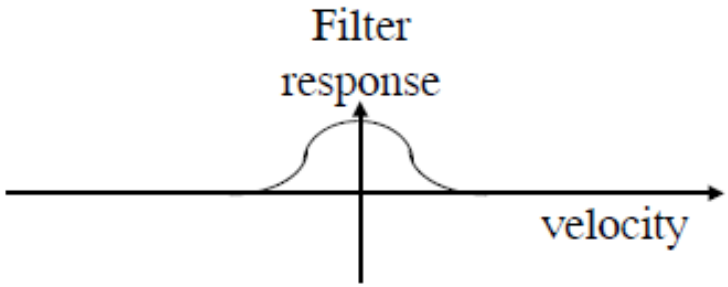
Orientation in visual space-time

- Velocity
  - Consider the response to a horizontally moving pattern of ***r***, ***l*** and ***s*** filters.



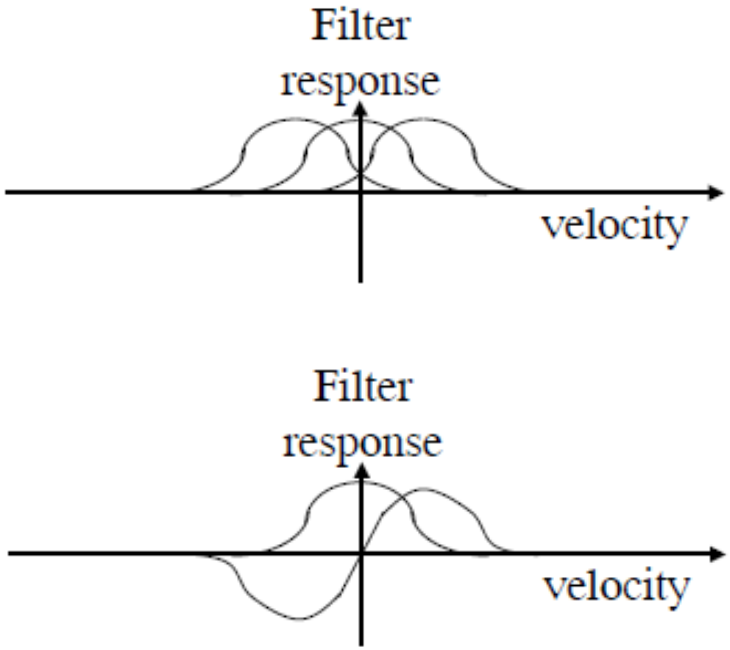
Orientation in visual space-time

- Velocity
  - Consider the response to a horizontally moving pattern of ***r***, ***l*** and ***s*** filters.



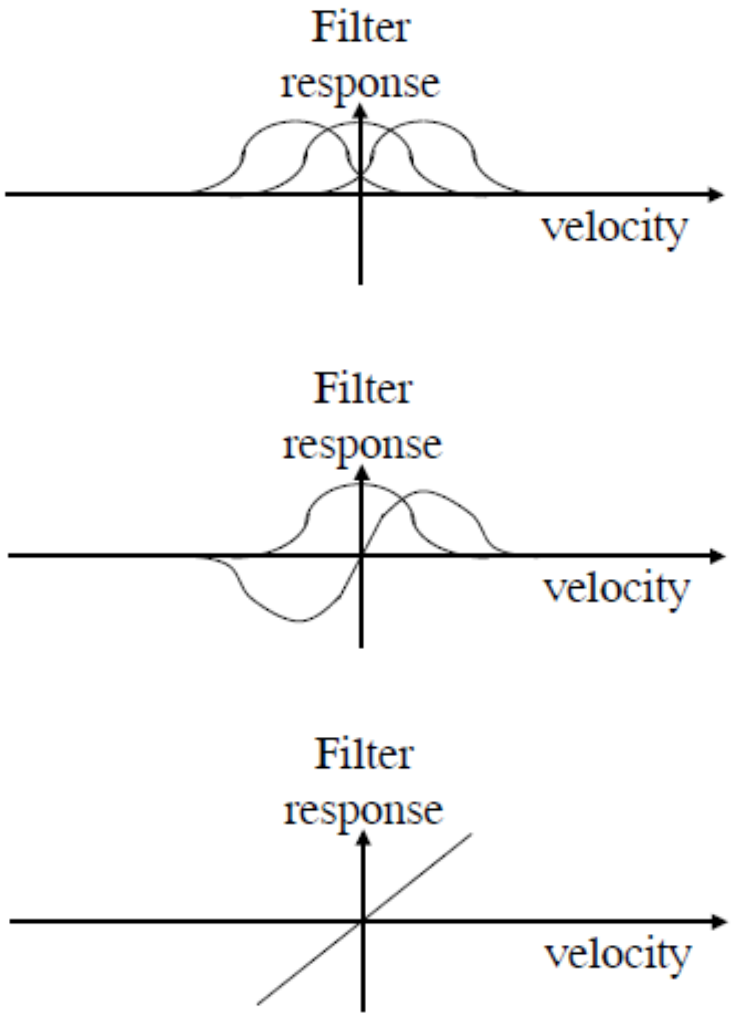
Orientation in visual space-time

- Velocity
  - Consider the response to a horizontally moving pattern of ***r***, ***l*** and ***s*** filters.
  - By taking the difference ***r-l*** we get a single response that is properly signed WRT velocity.



Orientation in visual space-time

- Velocity
  - Consider the response to a horizontally moving pattern of *r*, *l* and *s* filters.
  - By taking the difference *r-l* we get a single response that is properly signed WRT velocity.
  - Dividing to get *(r-l)/s* yields a response that is (approximately) linear with velocity.



Orientation in visual space-time

- Velocity: Comparison with the optical flow constraint equation

$$E_x u + E_y v + E_t = 0$$

Orientation in visual space-time

- Velocity: Comparison with the optical flow constraint equation

$$E_x u + E_y v + E_t = 0$$

- Let us restrict consideration to one spatial dimension + time

$$E_x u + E_t = 0$$

- Now, we can directly solve for (1D) velocity

$$u = -E_t / E_x$$

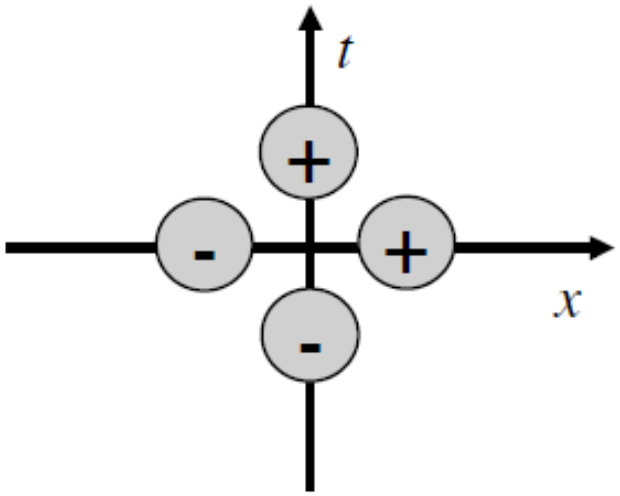
Orientation in visual space-time

- Velocity: Comparison with the optical flow constraint equation
  - In contrast, let us consider simple differential filters for leftward and rightward movement

Orientation in visual space-time

- Velocity: Comparison with the optical flow constraint equation
  - In contrast, let us consider simple differential filters for leftward and rightward movement

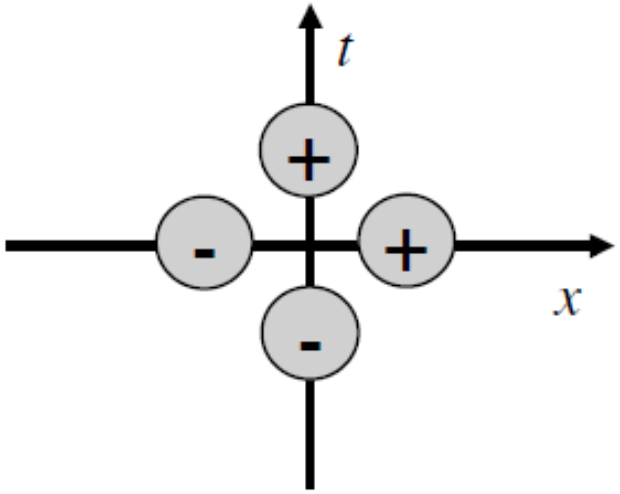
$$\tilde{l} = (E_t + E_x)/2$$



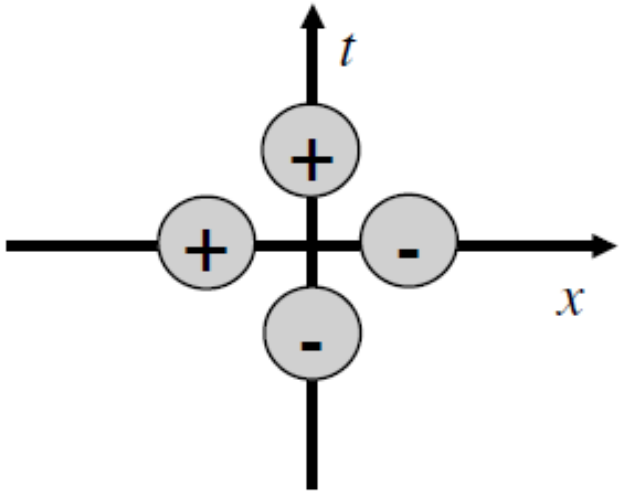
Orientation in visual space-time

- Velocity: Comparison with the optical flow constraint equation
  - In contrast, let us consider simple differential filters for leftward and rightward movement

$$\tilde{l} = (E_t + E_x)/2$$



$$\tilde{r} = (E_t - E_x)/2$$



Orientation in visual space-time

- Velocity: Comparison with the optical flow constraint equation
  - In contrast, let us consider simple differential filters for leftward and rightward movement

$$\tilde{l} = (E_t + E_x)/2 \qquad \tilde{r} = (E_t - E_x)/2$$

Orientation in visual space-time

- Velocity: Comparison with the optical flow constraint equation
  - In contrast, let us consider simple differential filters for leftward and rightward movement

$$\tilde{l} = (E_t + E_x)/2 \qquad \tilde{r} = (E_t - E_x)/2$$

  - Indeed, as we are concerned not with sign for a given direction, but rather magnitude, it suffices to consider the squared filter responses

$$l = (E_t^2 + 2E_xE_t + E_x^2)/4 \qquad r = (E_t^2 - 2E_xE_t + E_x^2)/4$$

Orientation in visual space-time

- Velocity: Comparison with the optical flow constraint equation

$$l = (E_t^2 + 2E_xE_t + E_x^2)/4 \qquad r = (E_t^2 - 2E_xE_t + E_x^2)/4$$

Orientation in visual space-time

- Velocity: Comparison with the optical flow constraint equation

$$l = (E_t^2 + 2E_xE_t + E_x^2)/4 \qquad r = (E_t^2 - 2E_xE_t + E_x^2)/4$$

  - To get a single (signed) measure of movement, we difference the two squared responses

$$r - l = -E_xE_t$$

Orientation in visual space-time

- Velocity: Comparison with the optical flow constraint equation

$$l = (E_t^2 + 2E_xE_t + E_x^2)/4 \qquad r = (E_t^2 - 2E_xE_t + E_x^2)/4$$

  - To get a single (signed) measure of movement, we difference the two squared responses

$$r - l = -E_xE_t$$

  - Finally, to avoid being biased by locally large values of image contrast, we divide through by the square of a first-order measure of local contrast $s=E_x^2$

$$(r - l)/s = -E_t/E_x$$

Orientation in visual space-time

- Velocity: Comparison with the optical flow constraint equation

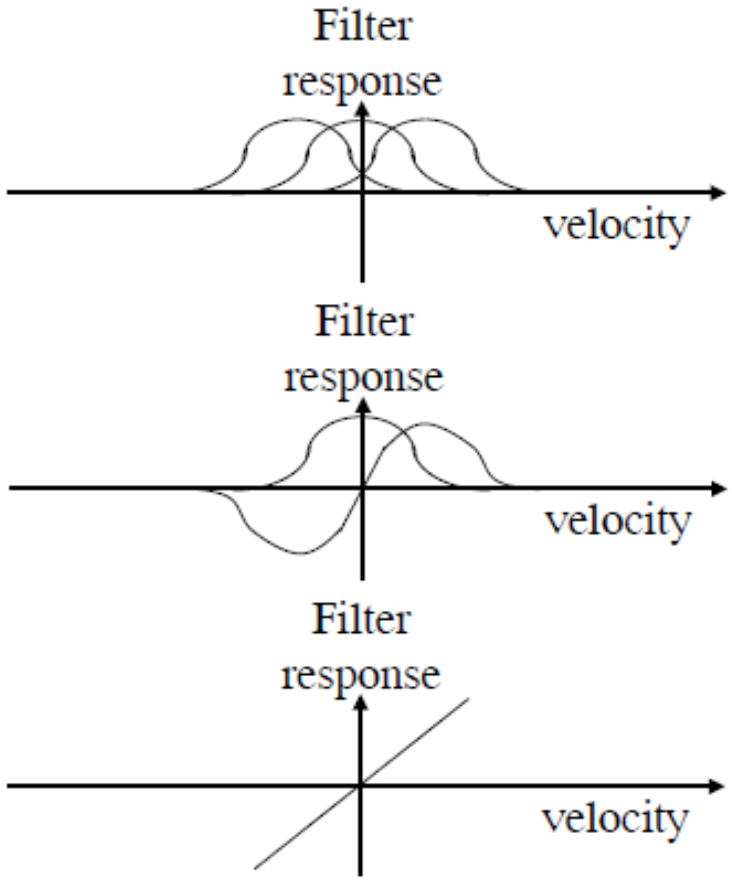
$$(r-l)/s = -E_t / E_x$$

  - We now recognize that this computation of velocity is equivalent to that based on the optical flow constraint equation

$$E_x u + E_t = 0 \Rightarrow u = -E_t / E_x$$

Orientation in visual space-time

- Velocity: Conclusion
  - Oriented filters in visual space-time support the recovery of image velocity, along a particular direction.
  - A “bank” of such filters can be used to span direction and provide an approach to optical flow estimation.



Outline

- Introduction
- Intuition
- Orientation in visual space-time
- A representation for spatiotemporal patterns
- Spatiotemporal boundaries
- Application to detection and tracking

A representation for spatiotemporal patterns

- Motivation
  - When confronted with spatiotemporal data, an intelligent system can be overwhelmed by sheer quantity.
  - An initial organization would be a key enabler for dealing effectively with data of this nature.
  - The organization should afford distinctions that can guide subsequent processing.
  - Distinctions that go beyond that of simple velocity.

A representation for spatiotemporal patterns

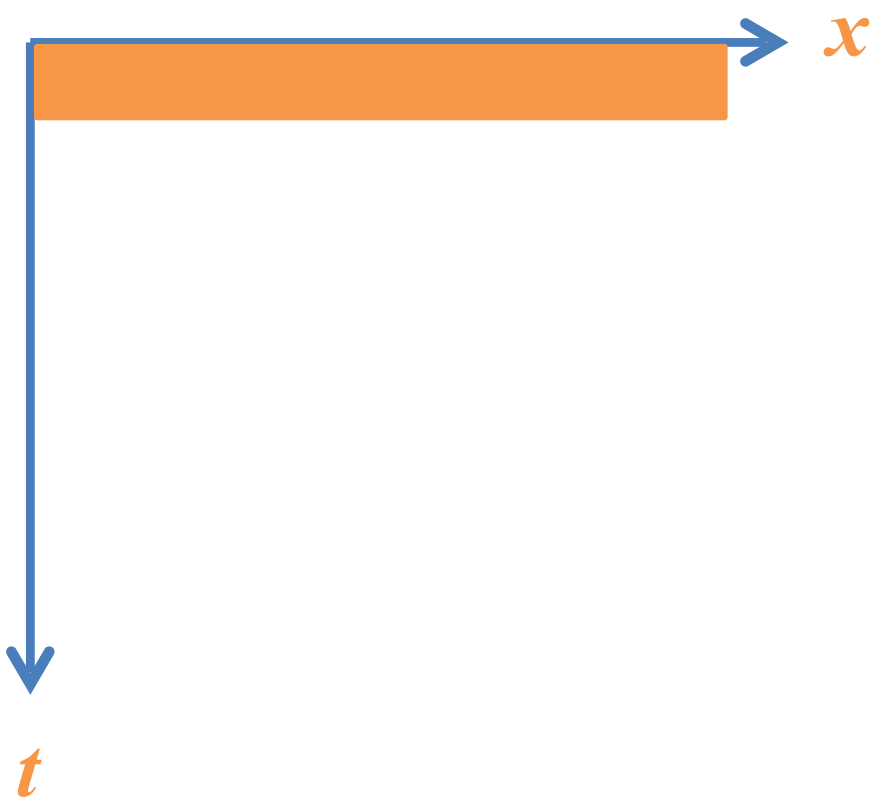
- General approach
  - Parse stream of spatiotemporal data into primitive, yet semantically meaningful categories at the earliest stages of processing.
  - Make distinctions along the following lines
    - Are the moving objects behaving coherently?
    - How much of the variance in the data is due to temporal brightness change?
    - Which portions of the data are simply too unstructured to support further analysis?

A representation for spatiotemporal patterns

- General approach
  - Integrate information across both the spatial and temporal dimensions.
  - Build on analysis of local orientation.
    - The simplest non-trivial characterization of local geometric structure.

A representation for spatiotemporal patterns

- Primitive patterns
  - Consider a spatiotemporal slice



A representation for spatiotemporal patterns

- Primitive patterns
  - Consider a spatiotemporal slice

A representation for spatiotemporal patterns

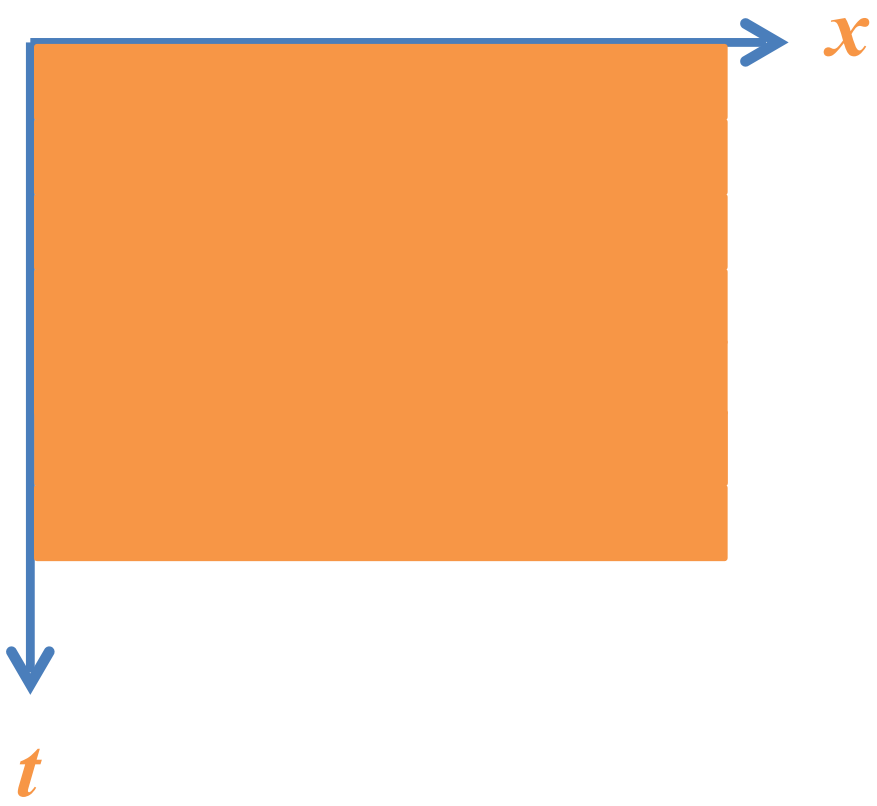
- Primitive patterns
  - Consider a spatiotemporal slice

A representation for spatiotemporal patterns

- Primitive patterns
  - Consider a spatiotemporal slice

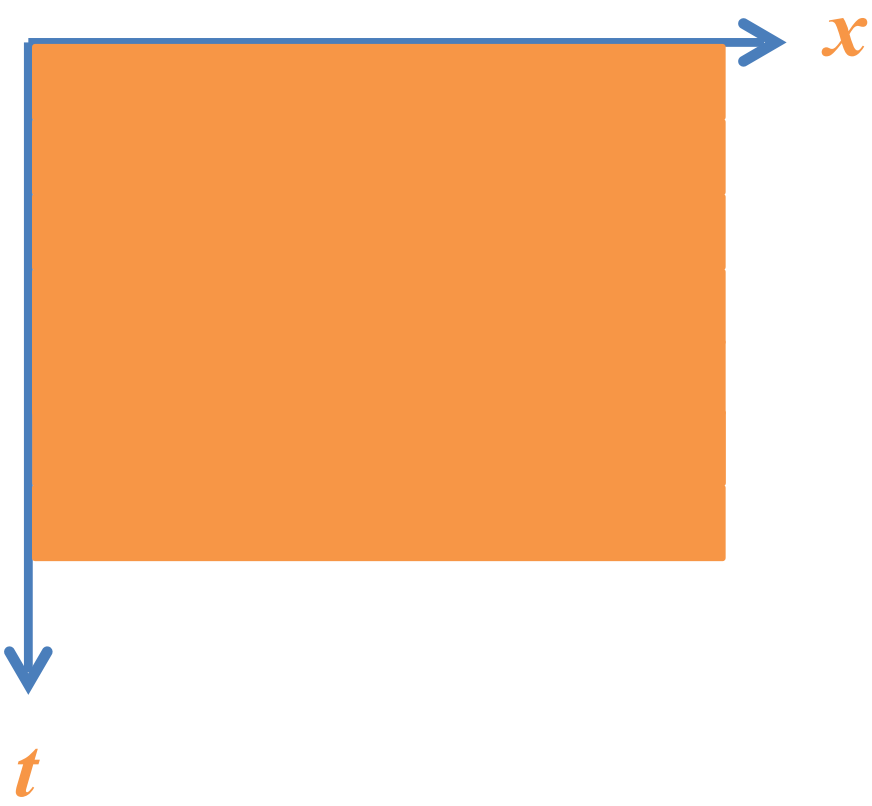
A representation for spatiotemporal patterns

- Primitive patterns
  - Consider a spatiotemporal slice



A representation for spatiotemporal patterns

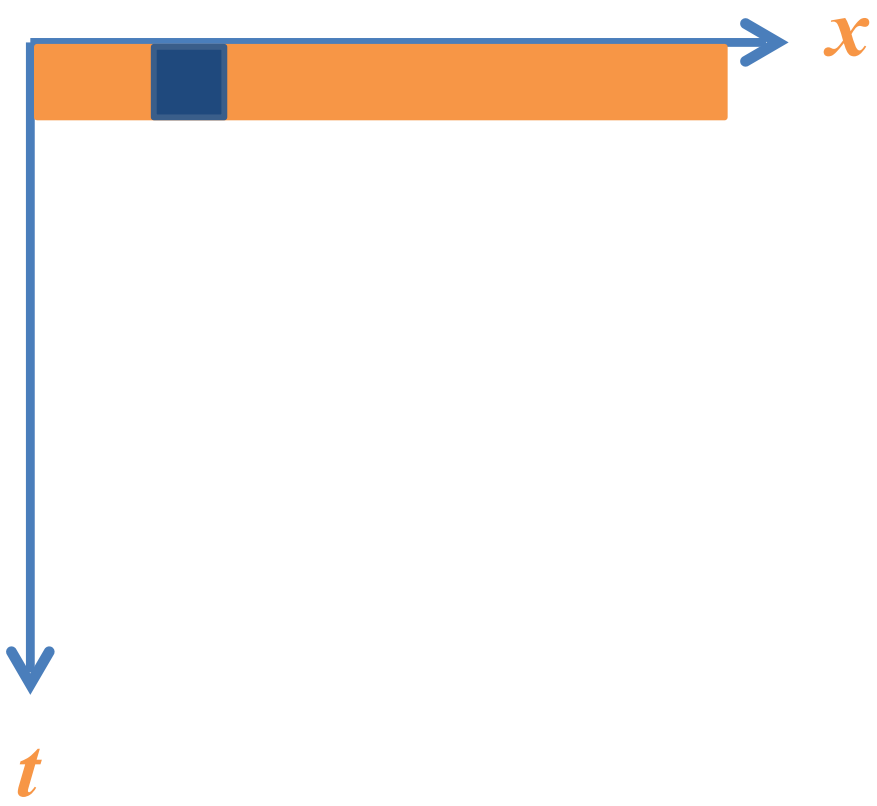
- Primitive patterns
  - Consider a spatiotemporal slice



This is a uniform pattern

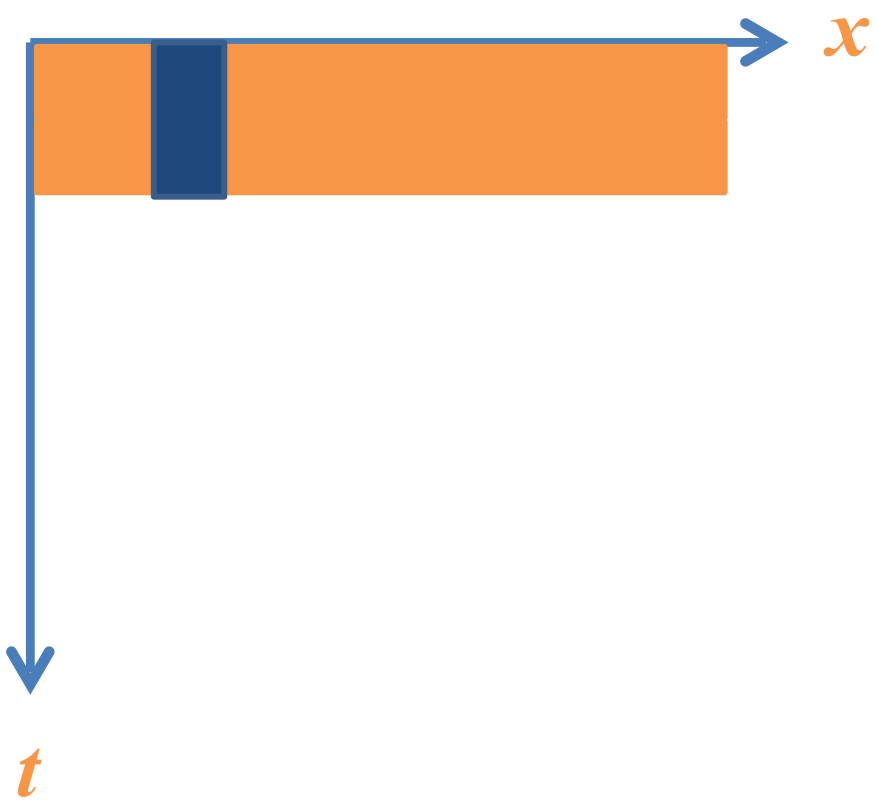
A representation for spatiotemporal patterns

- Primitive patterns
  - Consider a spatiotemporal slice



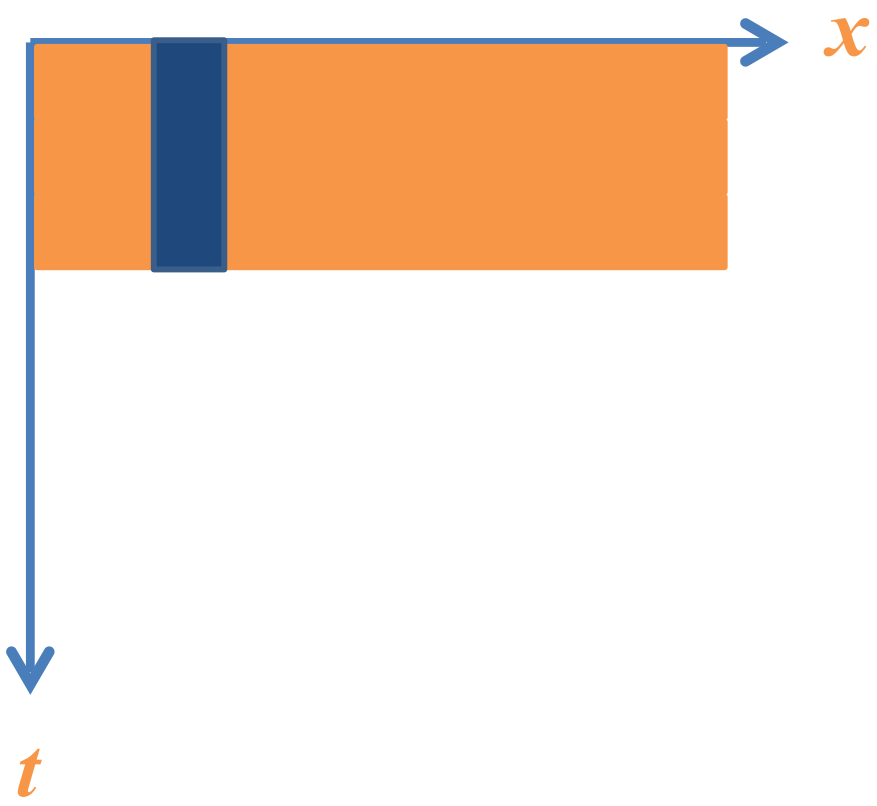
A representation for spatiotemporal patterns

- Primitive patterns
  - Consider a spatiotemporal slice



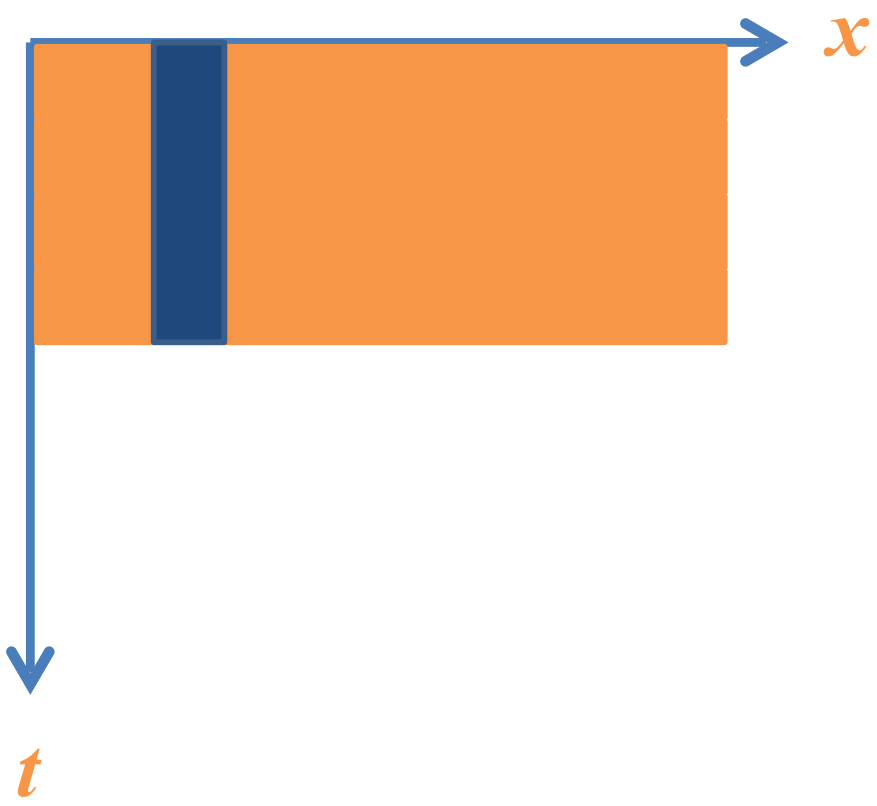
A representation for spatiotemporal patterns

- Primitive patterns
  - Consider a spatiotemporal slice



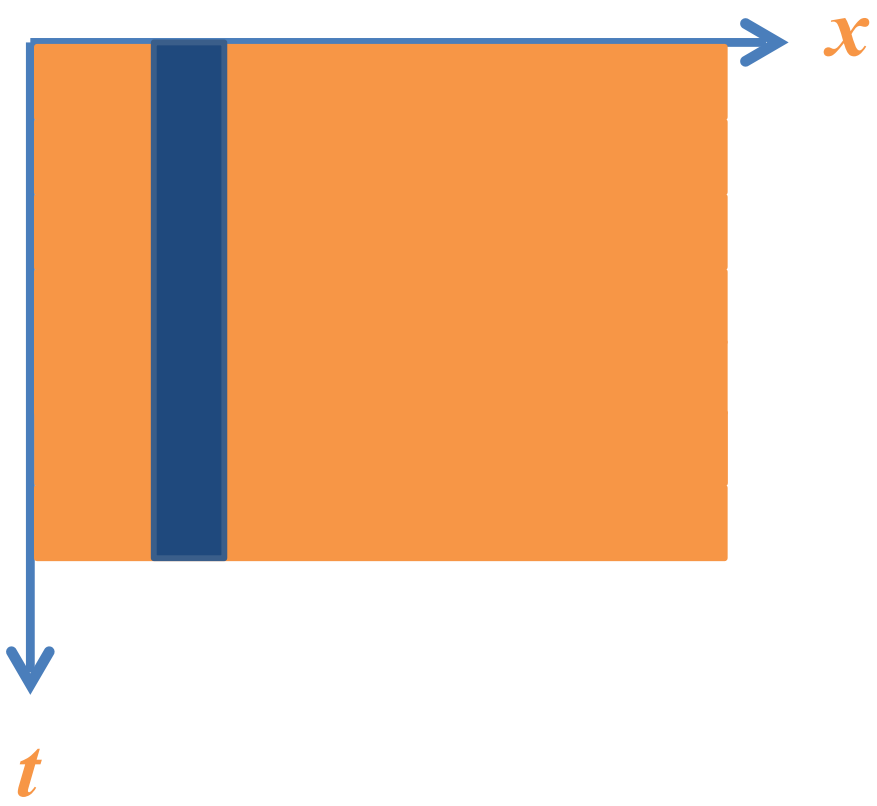
A representation for spatiotemporal patterns

- Primitive patterns
  - Consider a spatiotemporal slice



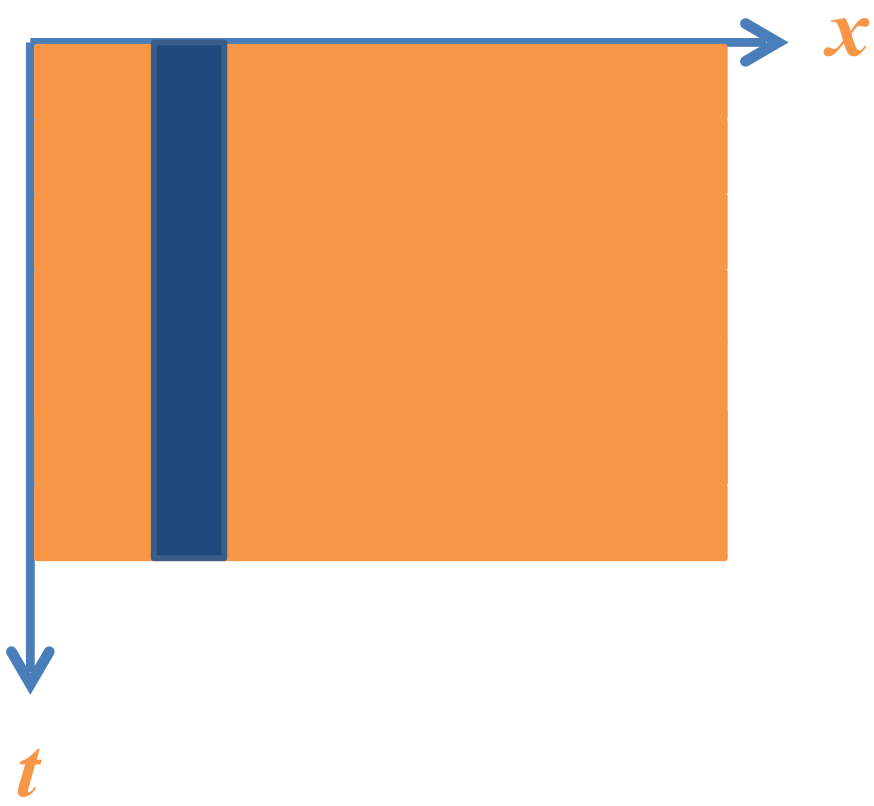
A representation for spatiotemporal patterns

- Primitive patterns
  - Consider a spatiotemporal slice



A representation for spatiotemporal patterns

- Primitive patterns
  - Consider a spatiotemporal slice



This is a
spatially structured,
temporally
static pattern

A representation for spatiotemporal patterns

- Primitive patterns
  - Consider a spatiotemporal slice

A representation for spatiotemporal patterns

- Primitive patterns
  - Consider a spatiotemporal slice

A representation for spatiotemporal patterns

- Primitive patterns
  - Consider a spatiotemporal slice

A representation for spatiotemporal patterns

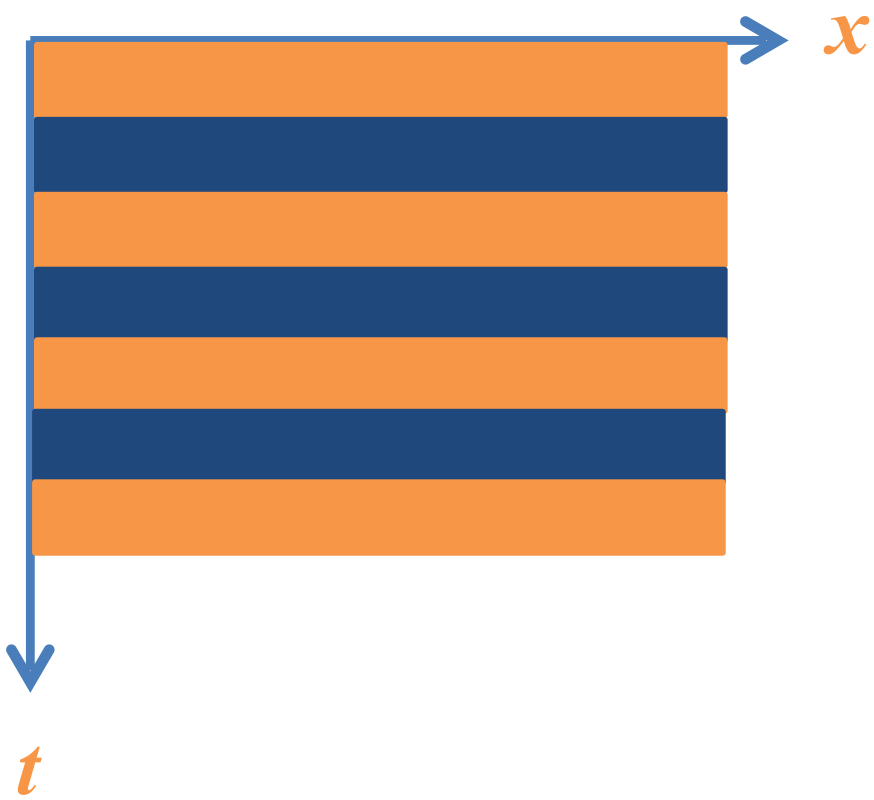
- Primitive patterns
  - Consider a spatiotemporal slice

A representation for spatiotemporal patterns

- Primitive patterns
  - Consider a spatiotemporal slice

A representation for spatiotemporal patterns

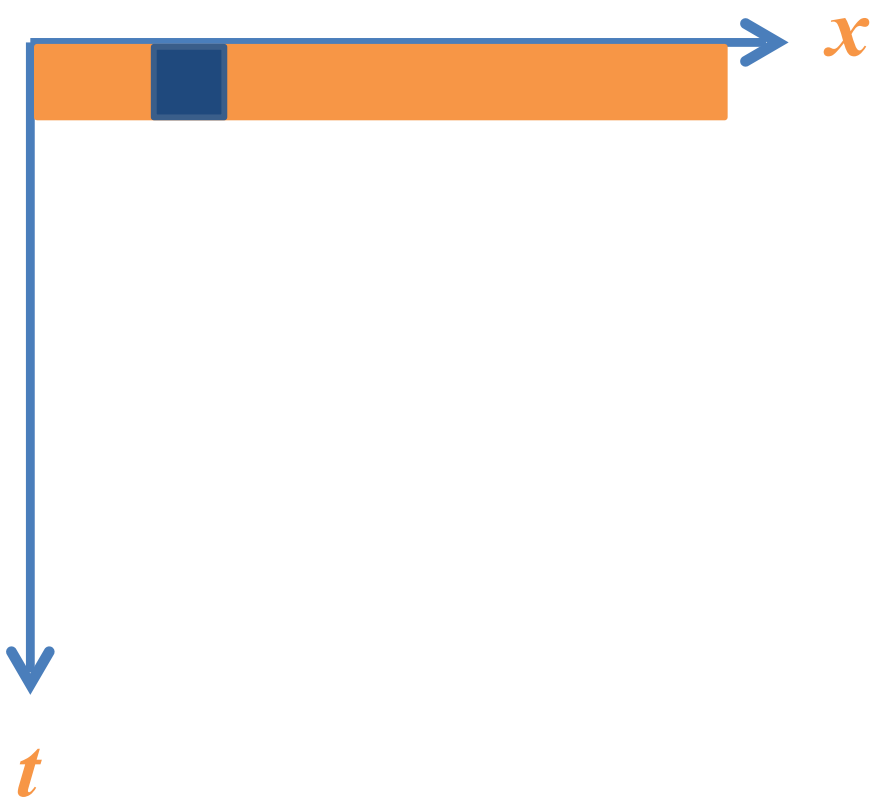
- Primitive patterns
  - Consider a spatiotemporal slice



This is a spatially uniform, temporally flickering pattern

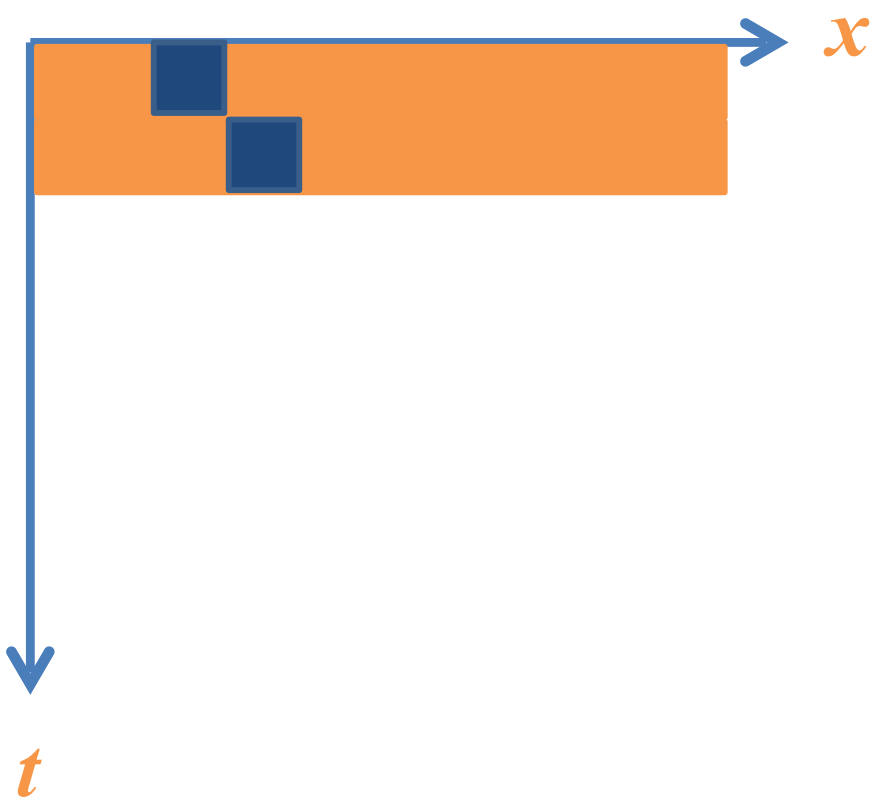
A representation for spatiotemporal patterns

- Primitive patterns
  - Consider a spatiotemporal slice



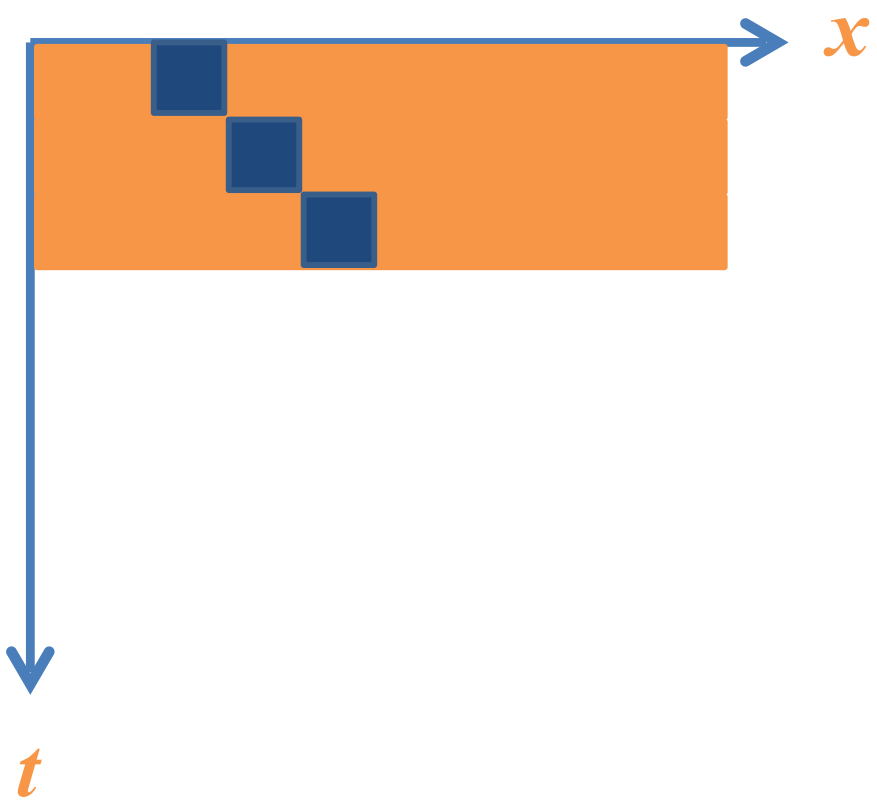
A representation for spatiotemporal patterns

- Primitive patterns
  - Consider a spatiotemporal slice



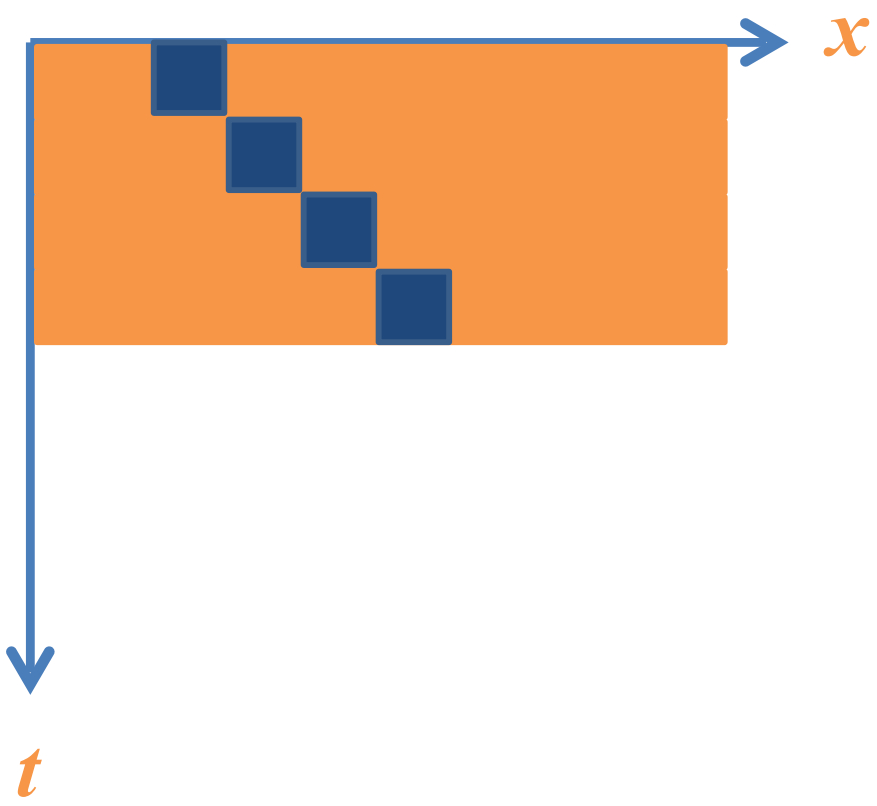
A representation for spatiotemporal patterns

- Primitive patterns
  - Consider a spatiotemporal slice



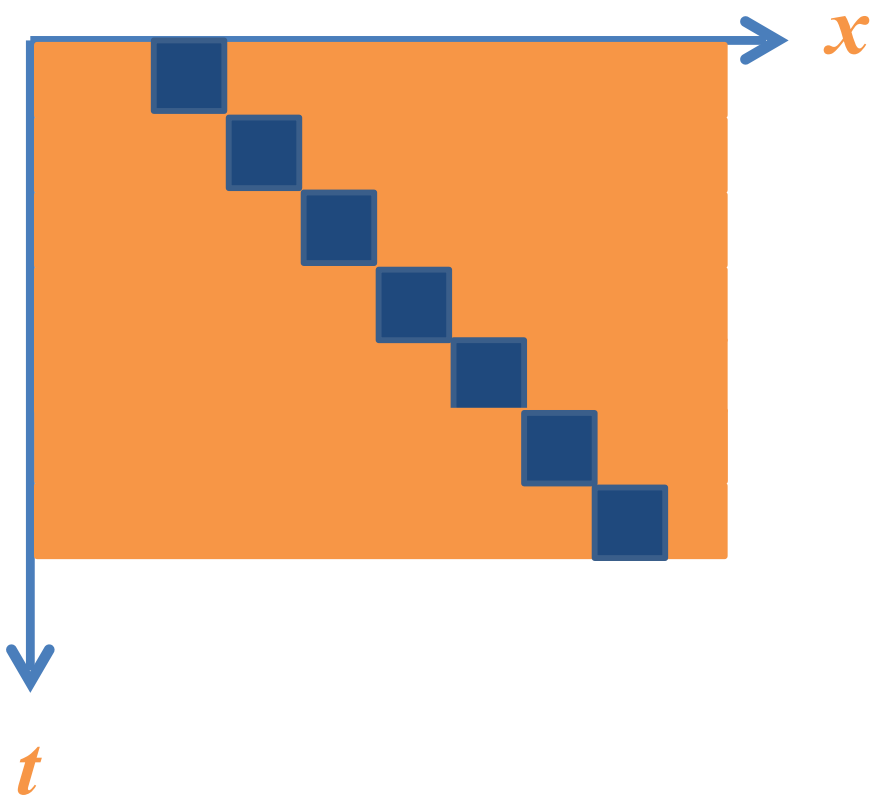
A representation for spatiotemporal patterns

- Primitive patterns
  - Consider a spatiotemporal slice



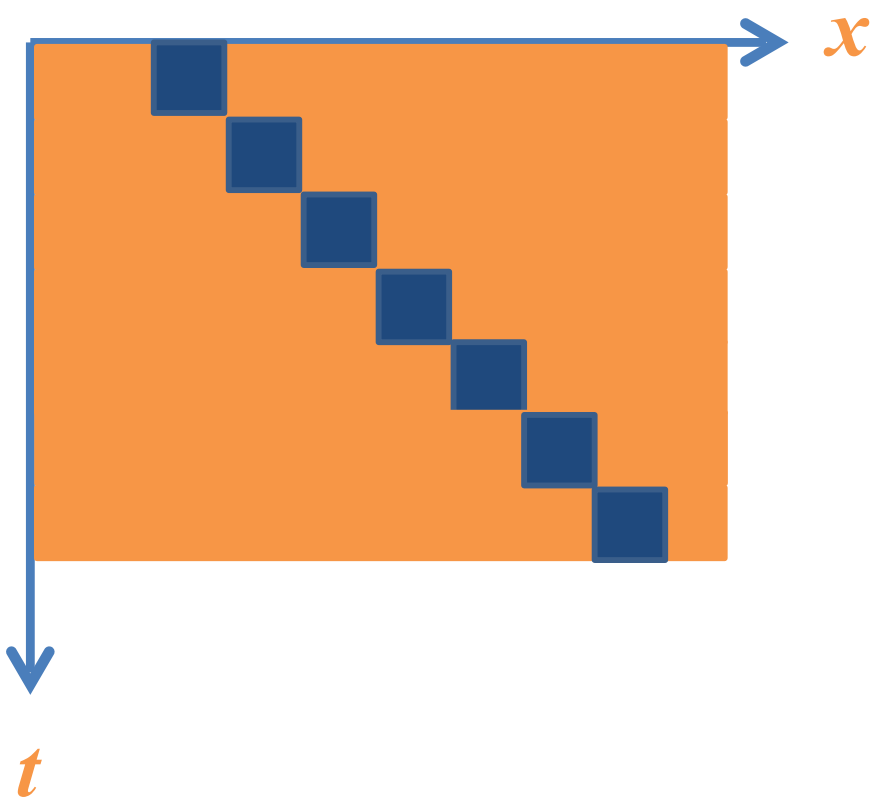
A representation for spatiotemporal patterns

- Primitive patterns
  - Consider a spatiotemporal slice



A representation for spatiotemporal patterns

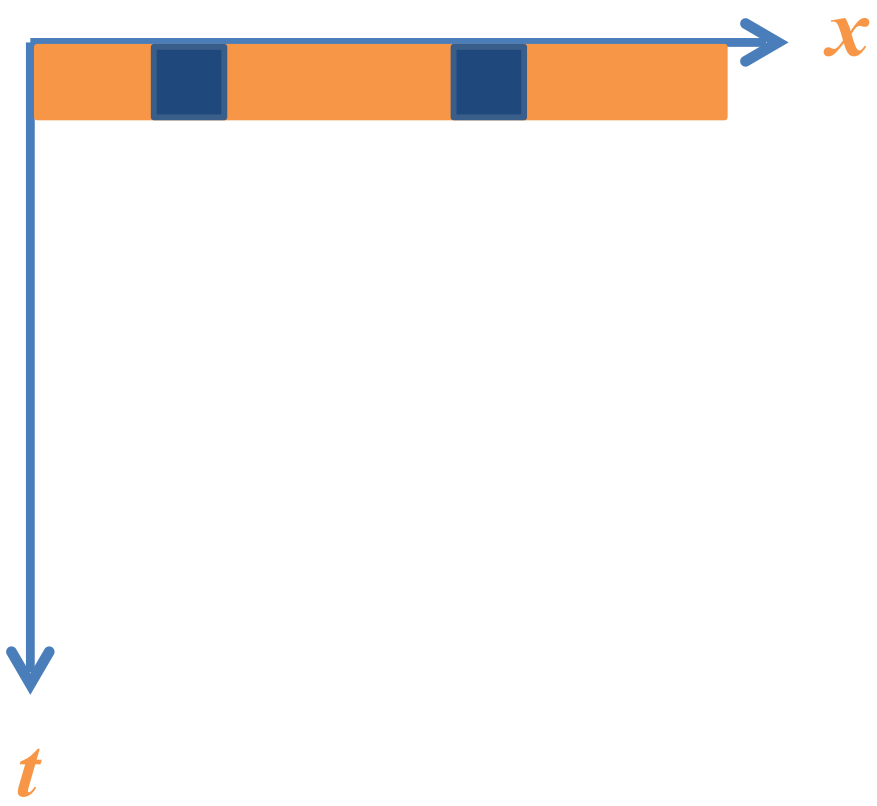
- Primitive patterns
  - Consider a spatiotemporal slice



This is a
spatially structured,
temporally
moving pattern

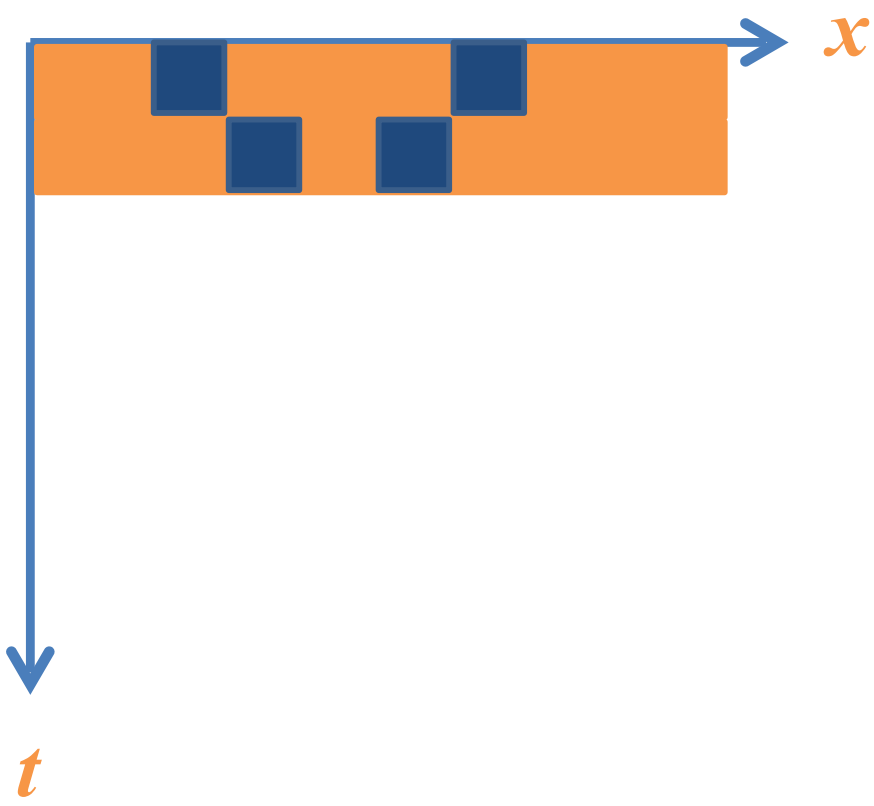
A representation for spatiotemporal patterns

- Primitive patterns
  - Consider a spatiotemporal slice



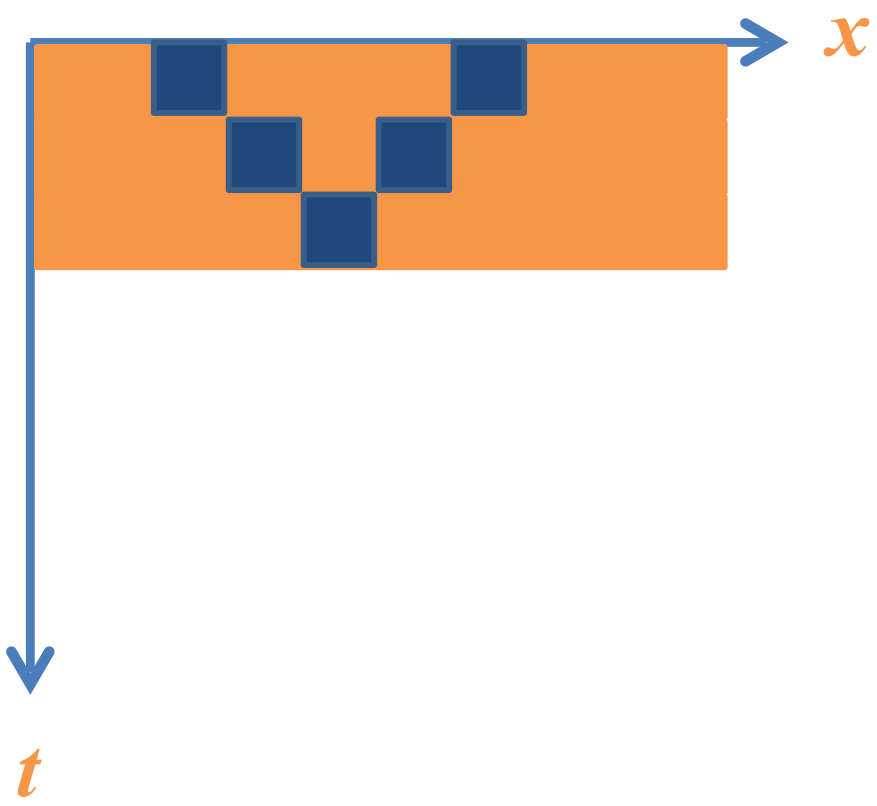
A representation for spatiotemporal patterns

- Primitive patterns
  - Consider a spatiotemporal slice



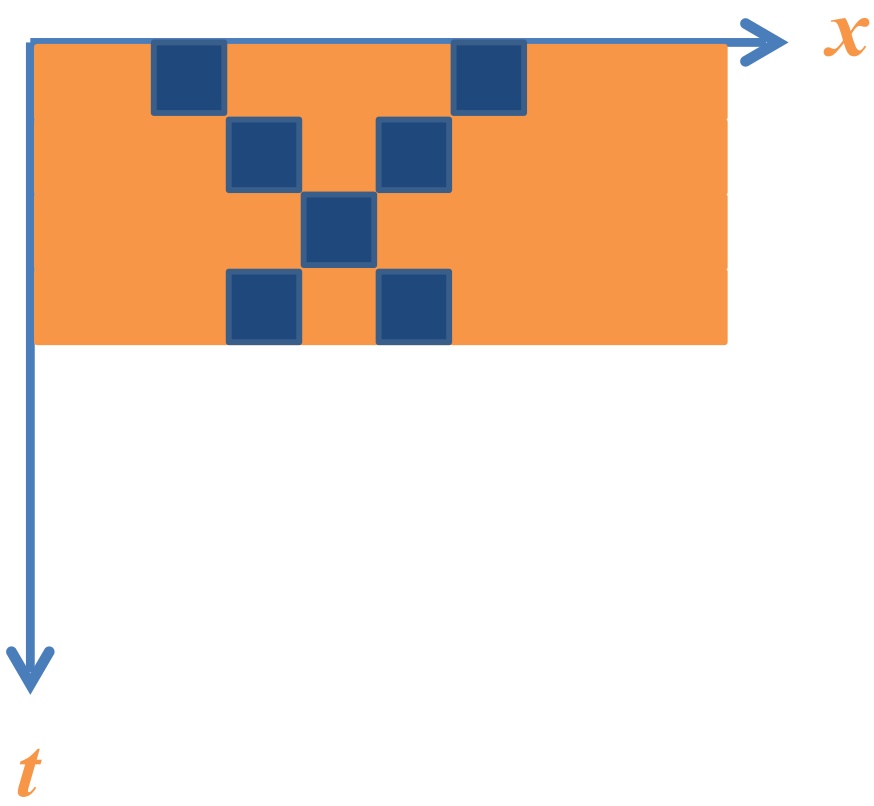
A representation for spatiotemporal patterns

- Primitive patterns
  - Consider a spatiotemporal slice



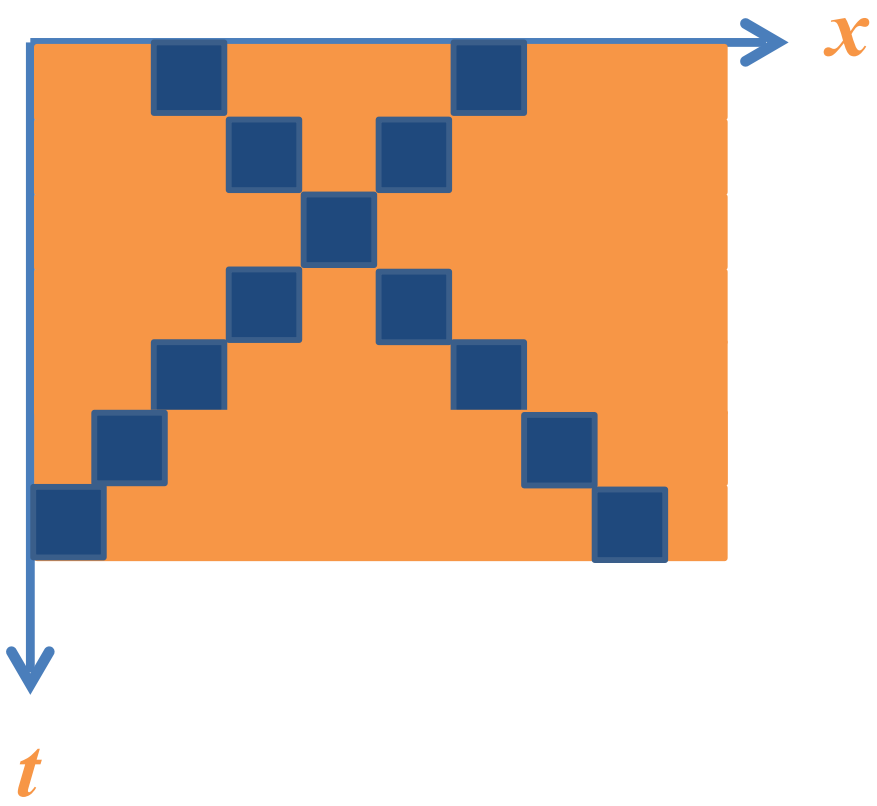
A representation for spatiotemporal patterns

- Primitive patterns
  - Consider a spatiotemporal slice



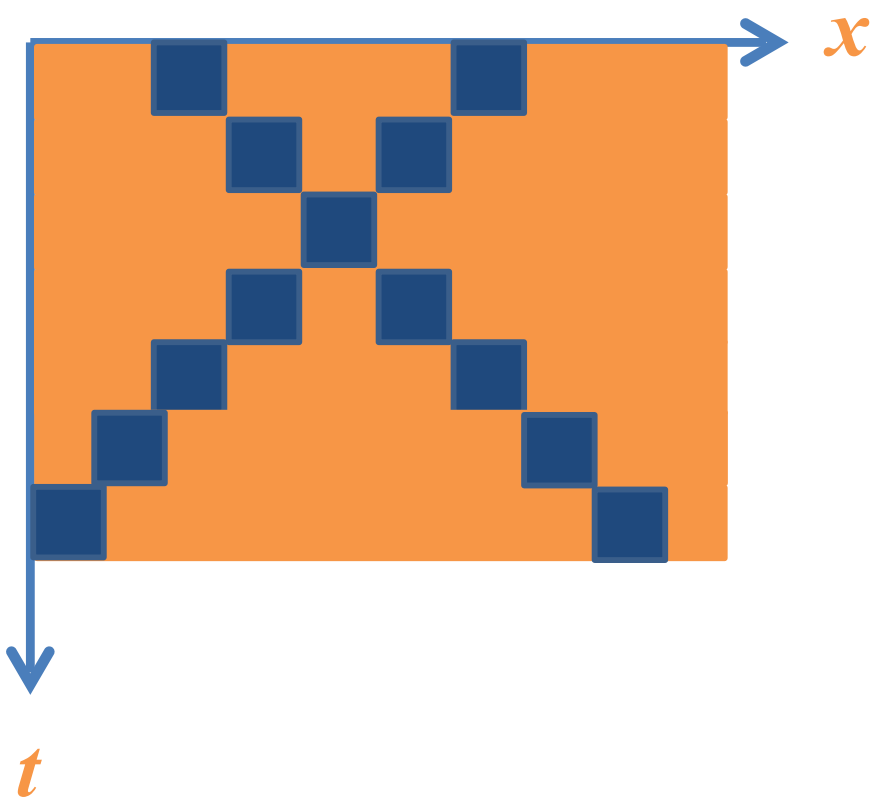
A representation for spatiotemporal patterns

- Primitive patterns
  - Consider a spatiotemporal slice



A representation for spatiotemporal patterns

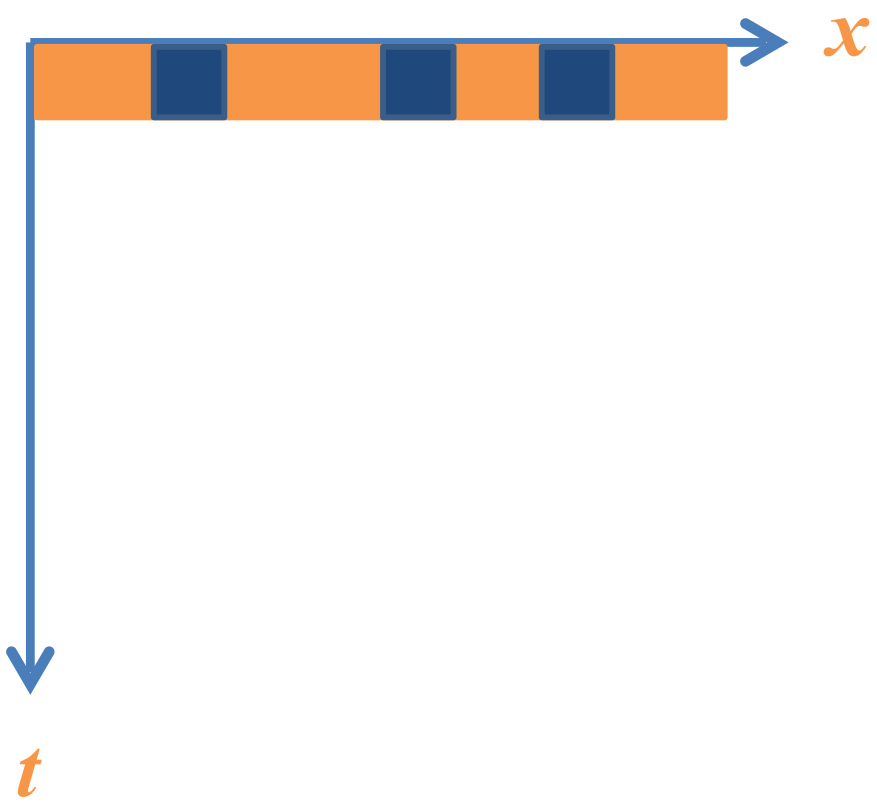
- Primitive patterns
  - Consider a spatiotemporal slice



This is a
spatially structured,
temporally
moving pattern
With > 1 motion

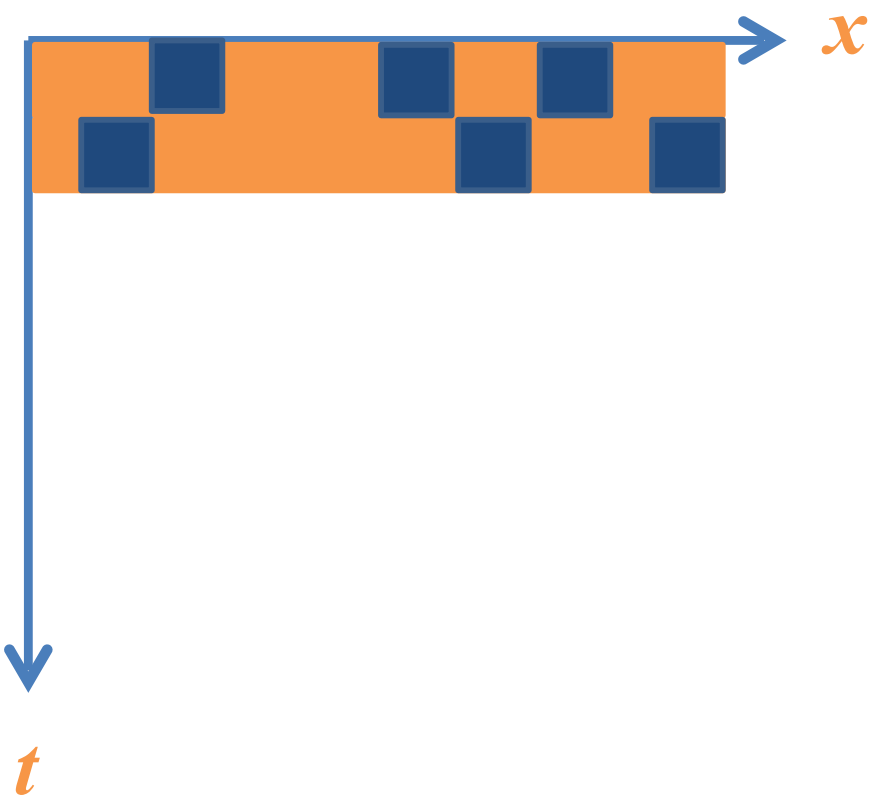
A representation for spatiotemporal patterns

- Primitive patterns
  - Consider a spatiotemporal slice



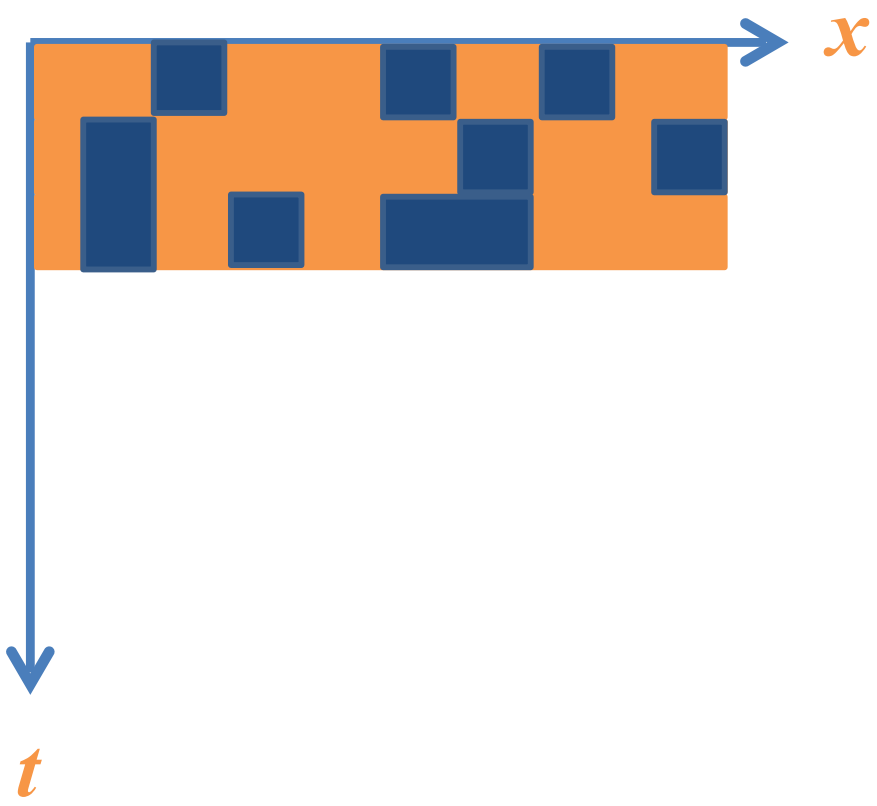
A representation for spatiotemporal patterns

- Primitive patterns
  - Consider a spatiotemporal slice



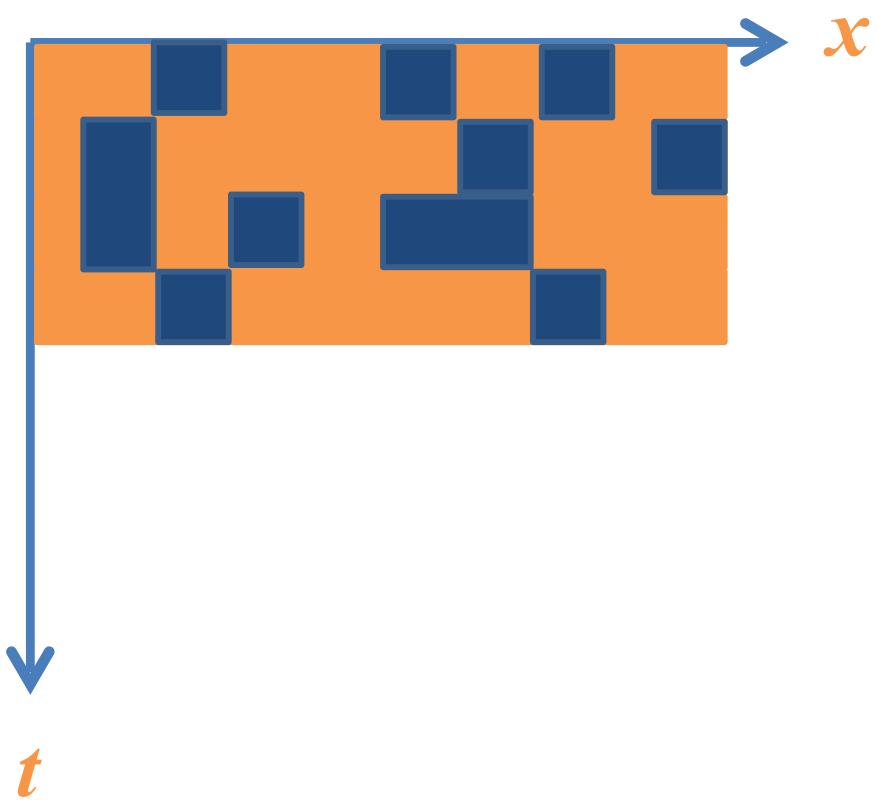
A representation for spatiotemporal patterns

- Primitive patterns
  - Consider a spatiotemporal slice



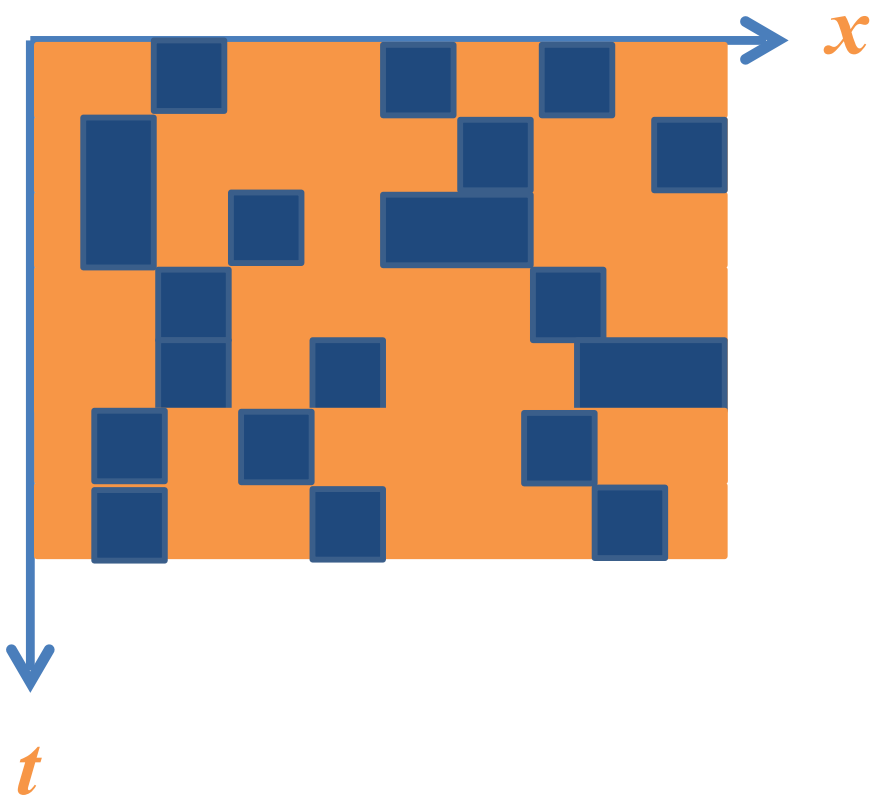
A representation for spatiotemporal patterns

- Primitive patterns
  - Consider a spatiotemporal slice



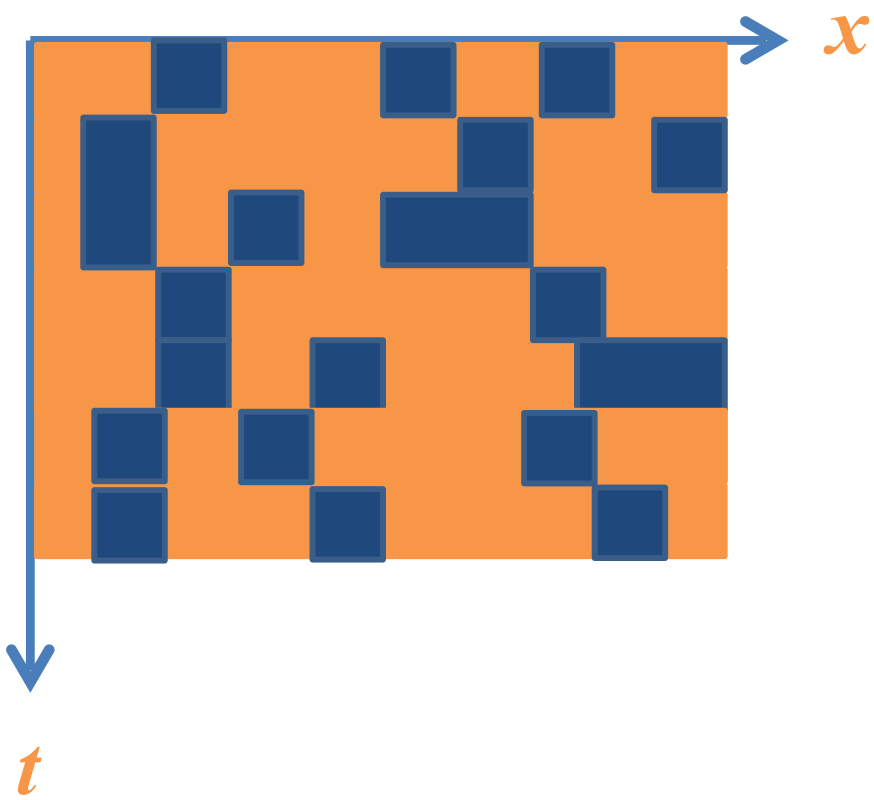
A representation for spatiotemporal patterns

- Primitive patterns
  - Consider a spatiotemporal slice



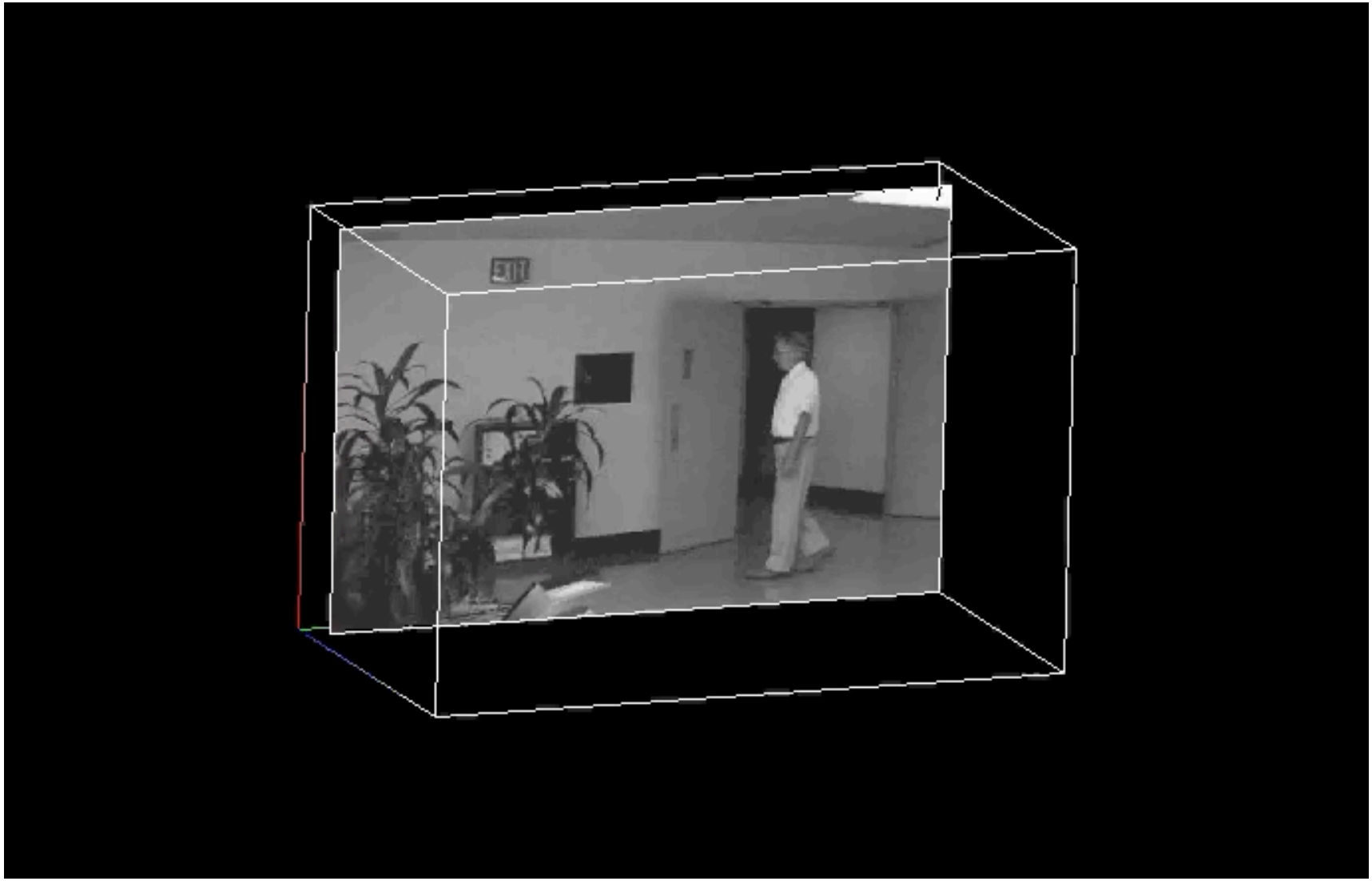
A representation for spatiotemporal patterns

- Primitive patterns
  - Consider a spatiotemporal slice

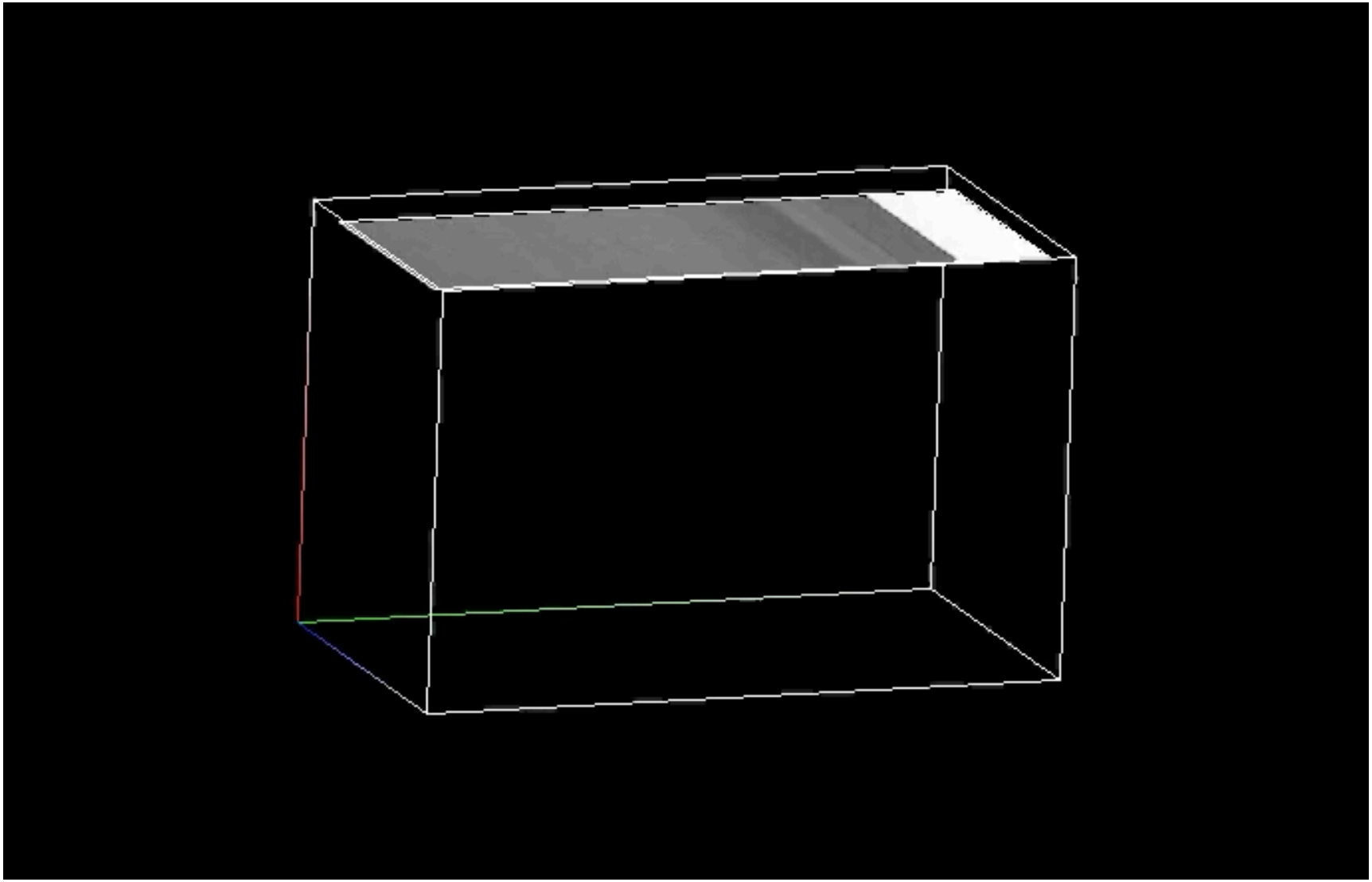


This is a
“randomly” structured,
in space and time

Example

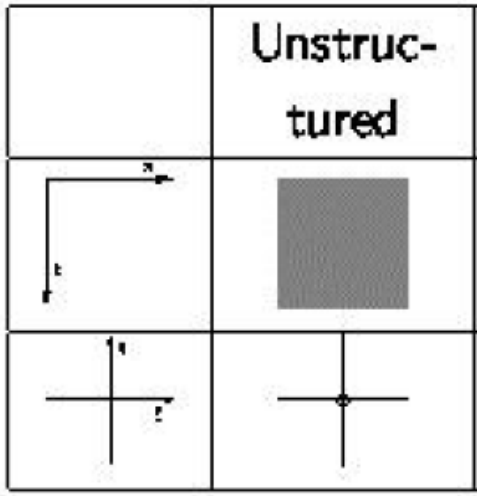


Example



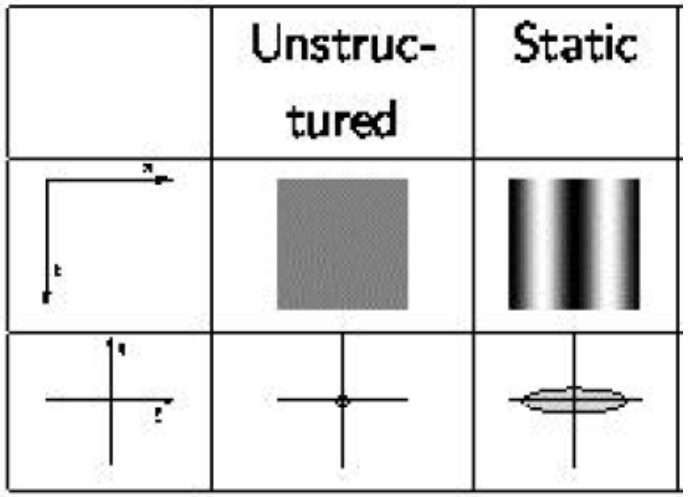
A representation for spatiotemporal patterns

- Primitives



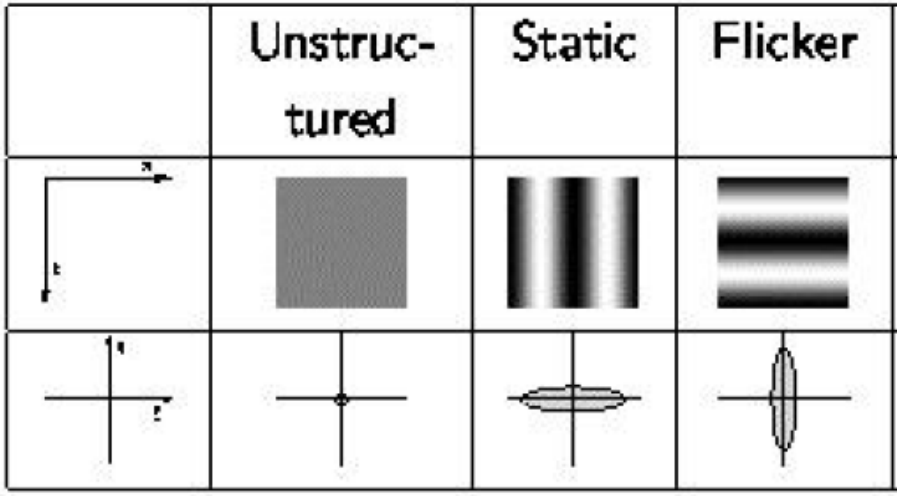
A representation for spatiotemporal patterns

- Primitives



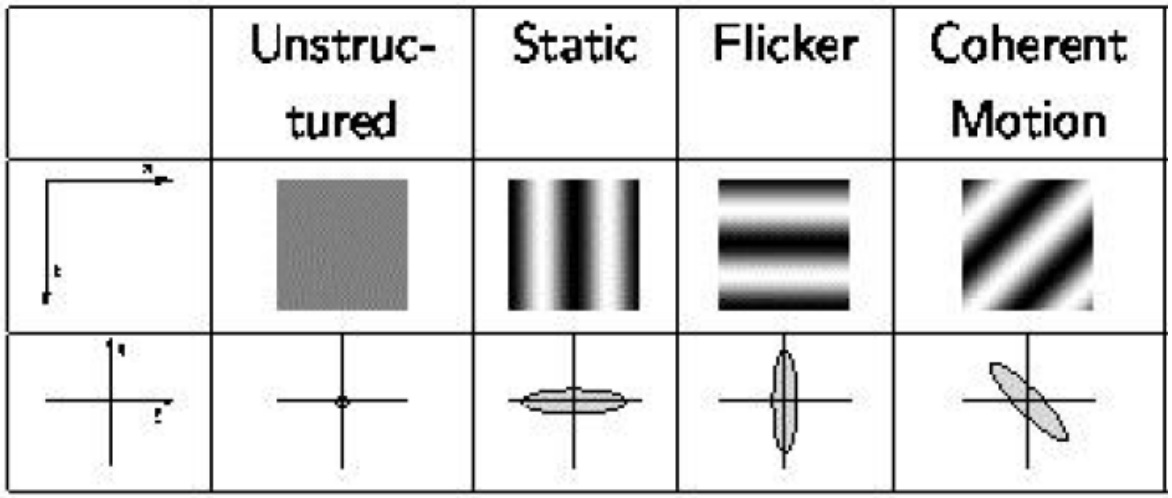
A representation for spatiotemporal patterns

- Primitives



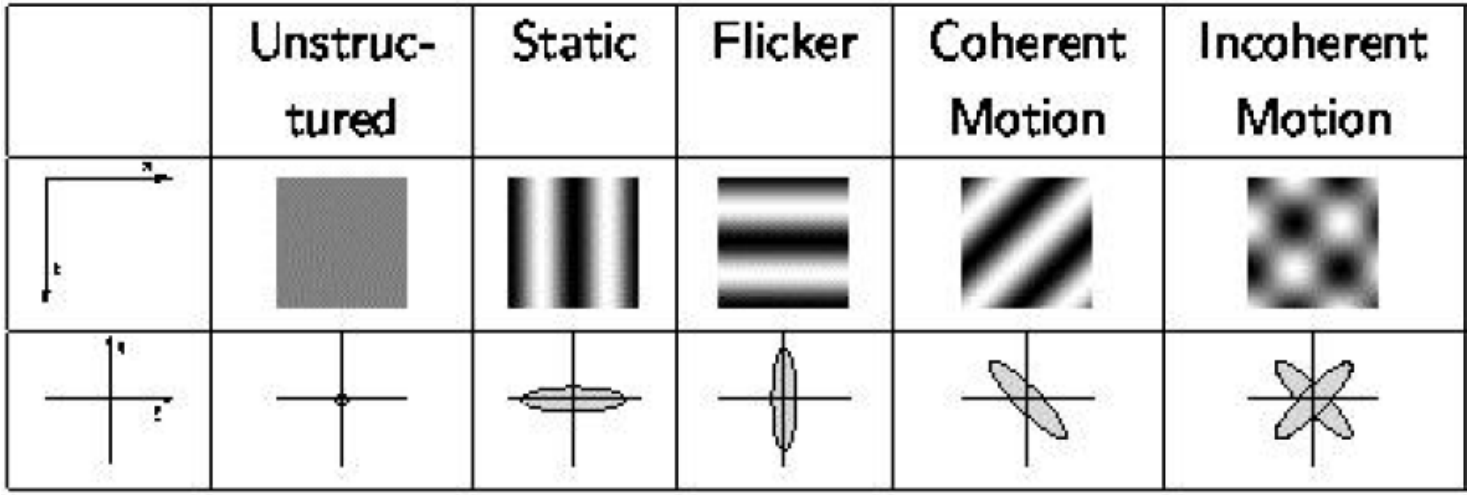
A representation for spatiotemporal patterns

- Primitives



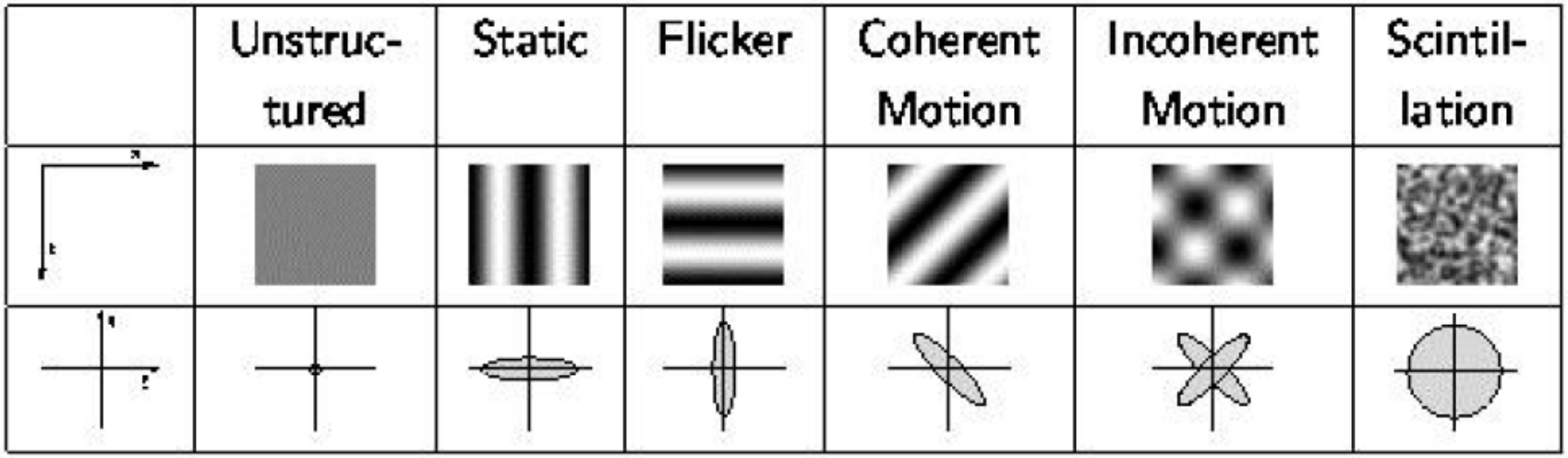
A representation for spatiotemporal patterns

- Primitives



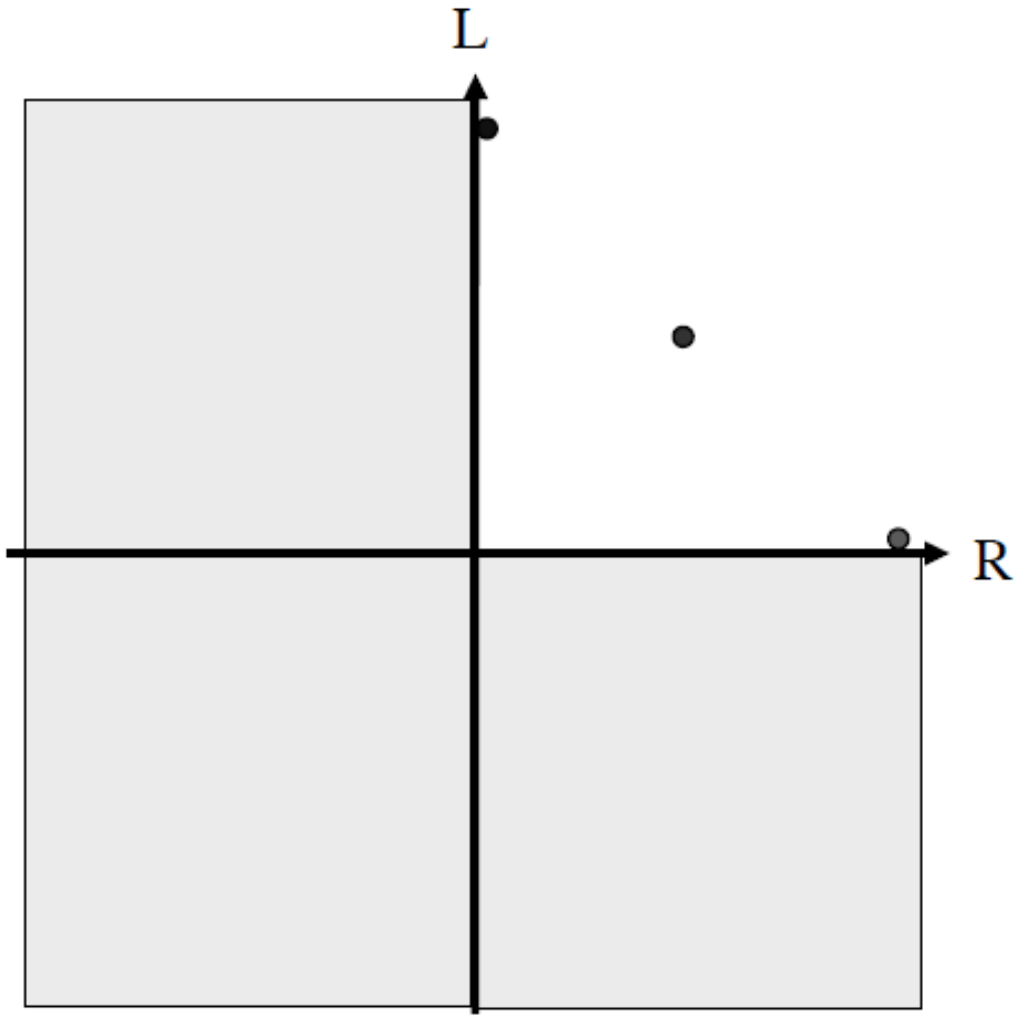
A representation for spatiotemporal patterns

- Primitives



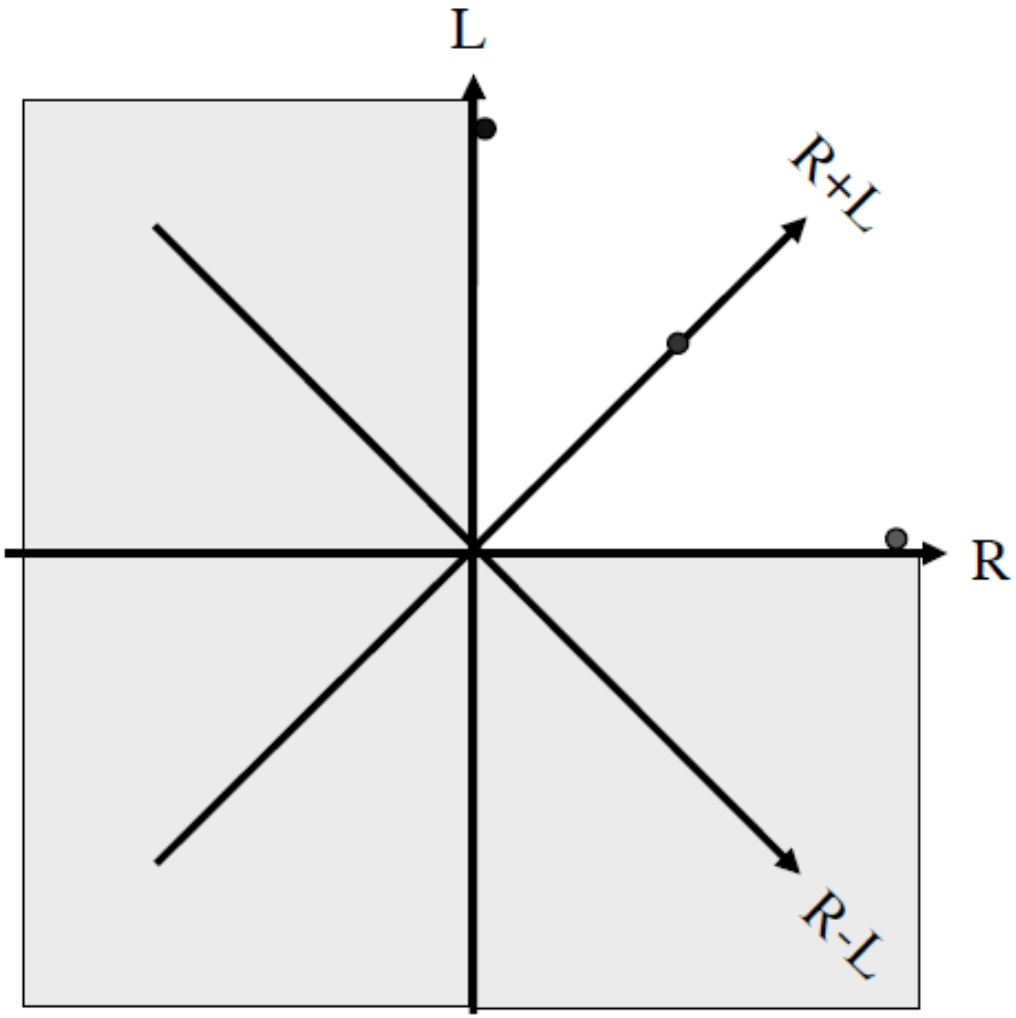
A representation for spatiotemporal patterns

- Opponency and summation
  - The R and L (U and D) components are ambiguous WRT coherent and incoherent motion.



A representation for spatiotemporal patterns

- Opponency and summation
  - The R and L (U and D) components are ambiguous WRT coherent and incoherent motion.
  - Solution: Combine via
    - opponency R-L (U-D)
    - summation R+L (U+D)
- Geometrically a rotation of coordinate axes.



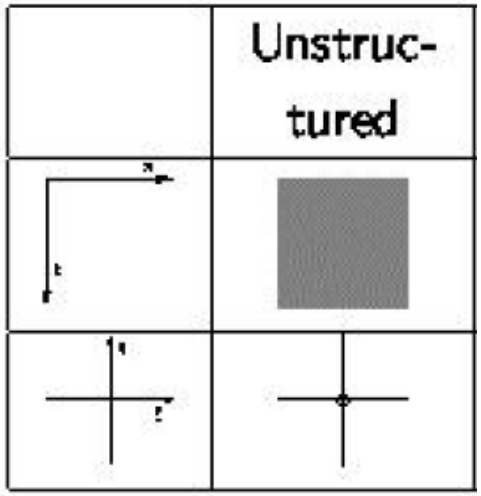
A representation for spatiotemporal patterns

- Proposal
  - A four band representation for both the x-t and y-t dimensions.

x-t	y-t
$\lvert R-L \rvert$	$\lvert U-D \rvert$
$R+L$	$U+D$
S_x	S_y
F_x	F_y

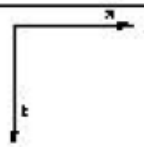
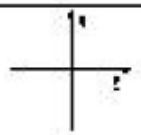
A representation for spatiotemporal patterns

- Primitives projected on representation



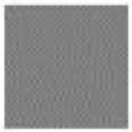
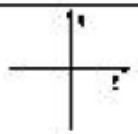
A representation for spatiotemporal patterns

- Primitives projected on representation

	Unstructured
	
	
$\|R - L\|$	0
$R + L$	0
S_x	0
F_x	0

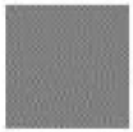
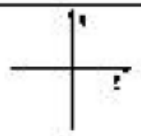
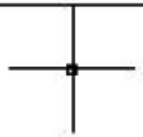
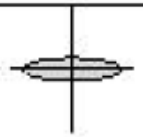
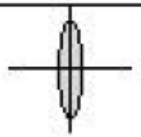
A representation for spatiotemporal patterns

- Primitives projected on representation

	Unstructured	Static
		
		
$\|R-L\|$	0	0
$R+L$	0	++
S_x	0	++
F_x	0	0

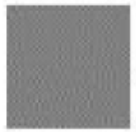
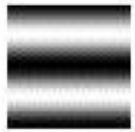
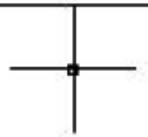
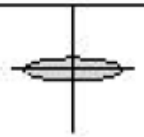
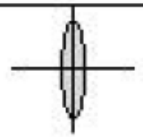
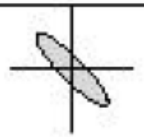
A representation for spatiotemporal patterns

- Primitives projected on representation

	Unstructured	Static	Flicker
			
				
$\|R-L\|$	0	0	0
$R+L$	0	++	++
S_x	0	++	0
F_x	0	0	++

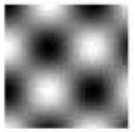
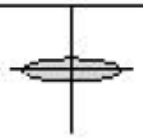
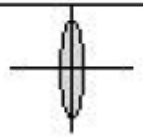
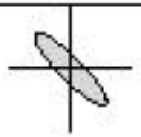
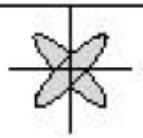
A representation for spatiotemporal patterns

- Primitives projected on representation

	Unstructured	Static	Flicker	Coherent Motion
				
				
$\lvert R-L\rvert$	0	0	0	++
$R+L$	0	++	++	++
S_x	0	++	0	+
F_x	0	0	++	+

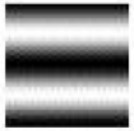
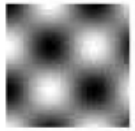
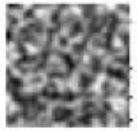
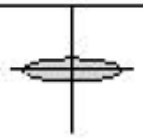
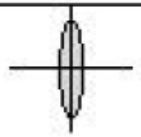
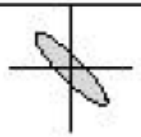
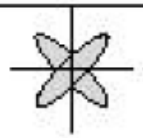
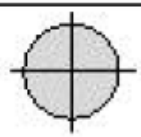
A representation for spatiotemporal patterns

- Primitives projected on representation

	Unstructured	Static	Flicker	Coherent Motion	Incoherent Motion
					
					
$\|R-L\|$	0	0	0	++	0
$R+L$	0	++	++	++	++++
S_x	0	++	0	+	+
F_x	0	0	++	+	+

A representation for spatiotemporal patterns

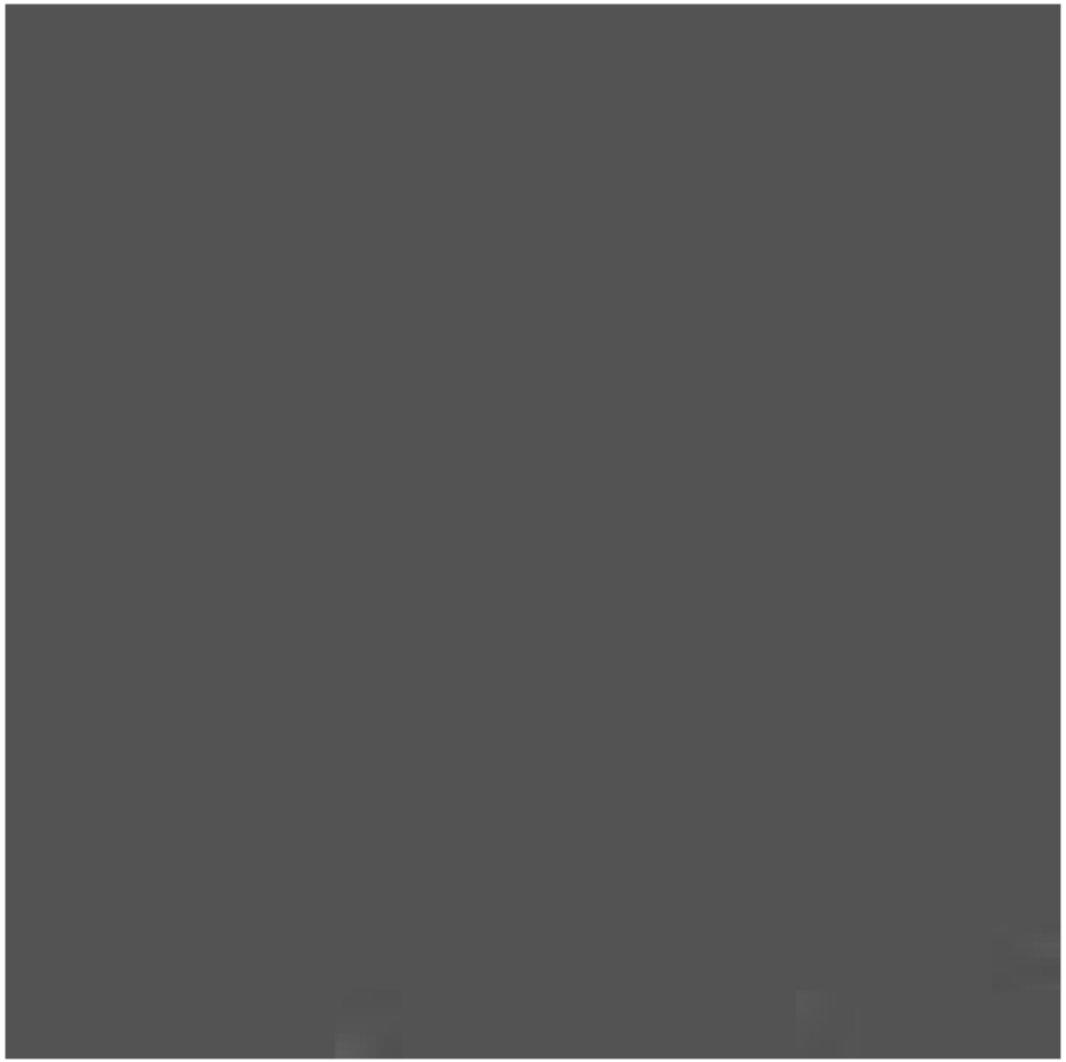
- Primitives projected on representation

	Unstructured	Static	Flicker	Coherent Motion	Incoherent Motion	Scintillation
						
						
$\|R-L\|$	0	0	0	++	0	0
$R+L$	0	++	++	++	+ + ++	++
S_x	0	++	0	+	+	+
F_x	0	0	++	+	+	+

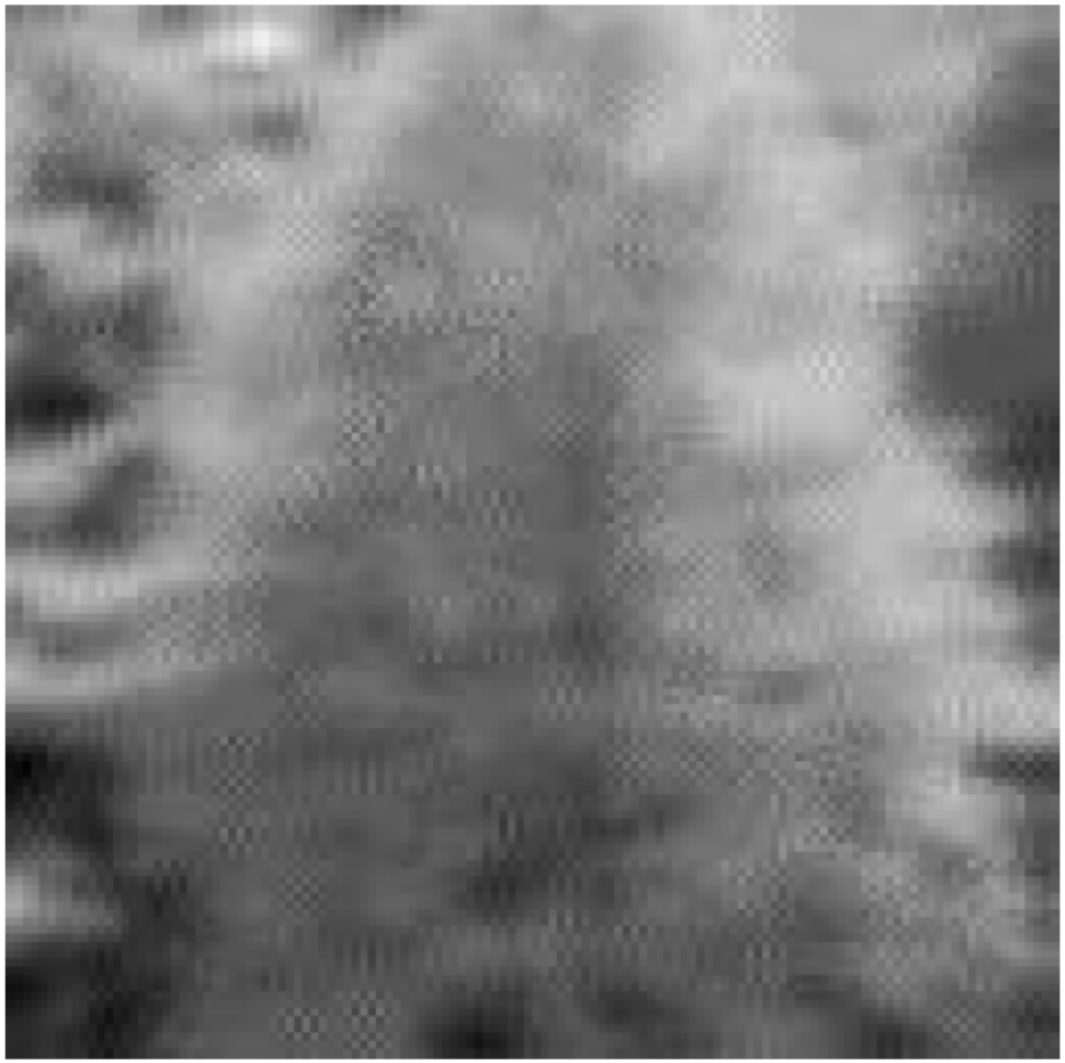
A representation for spatiotemporal patterns

- Examples (Wildes and Bergen, 2000)
  - A natural image sequence of each proposed class was acquired, (x,y,t) = (64,64,40).
    - Unstructured: featureless sky
    - static: motionless tree
    - flicker: smooth surface illuminated by lightning flashes
    - coherent motion: field of flowers under camera motion
    - incoherent motion: overlapped legs in complex motion
    - scintillation: rain striking a puddle
  - Each sequence brought under proposed representation.

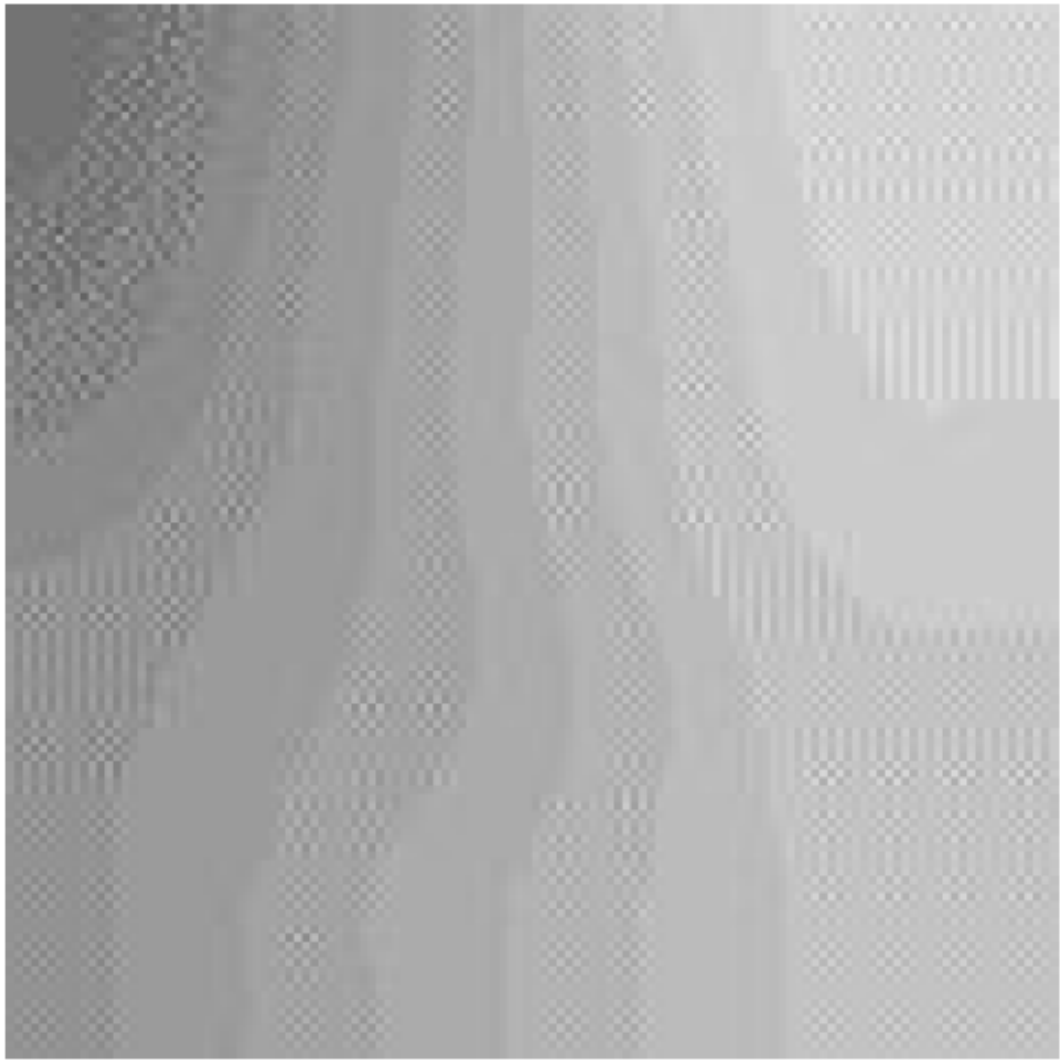
Unstructured



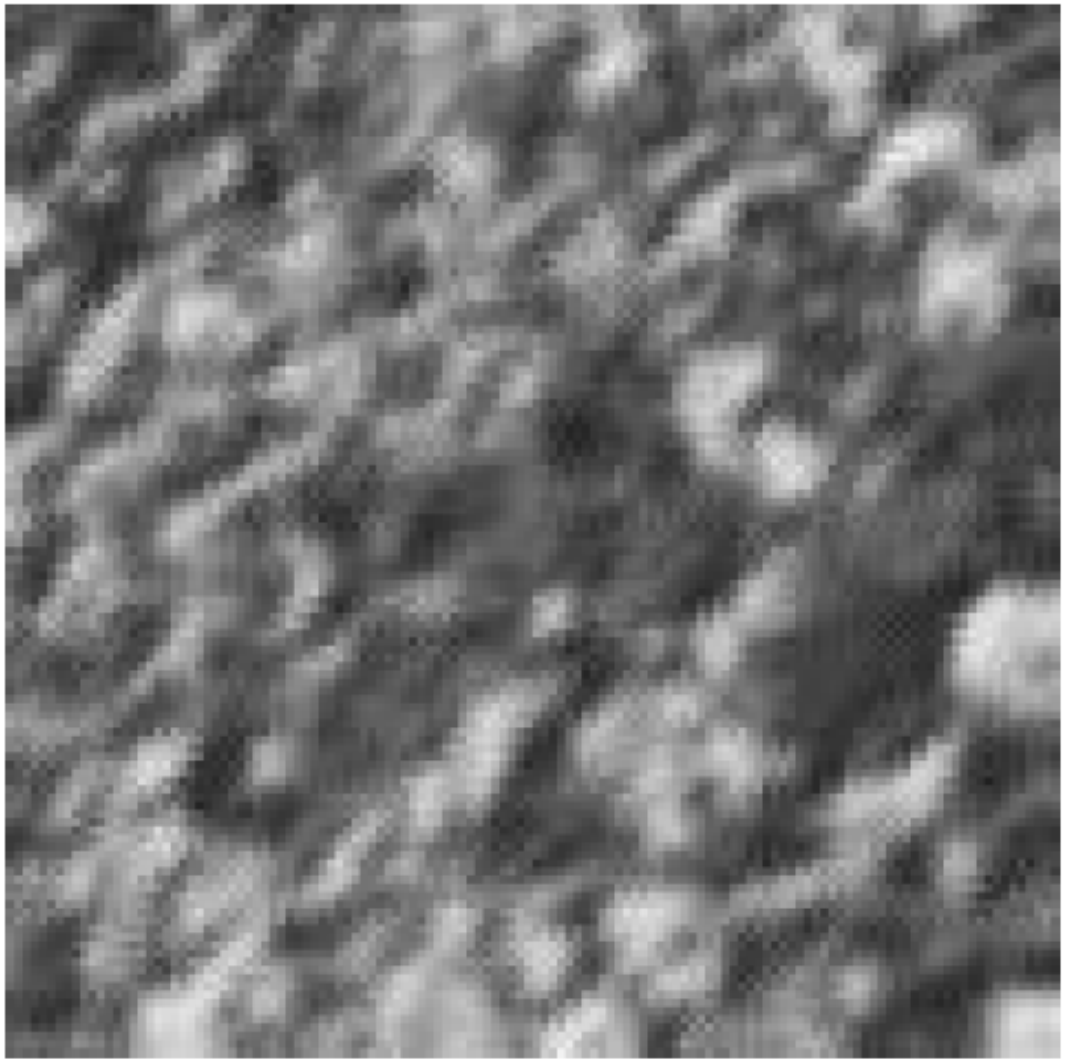
Static



Flicker

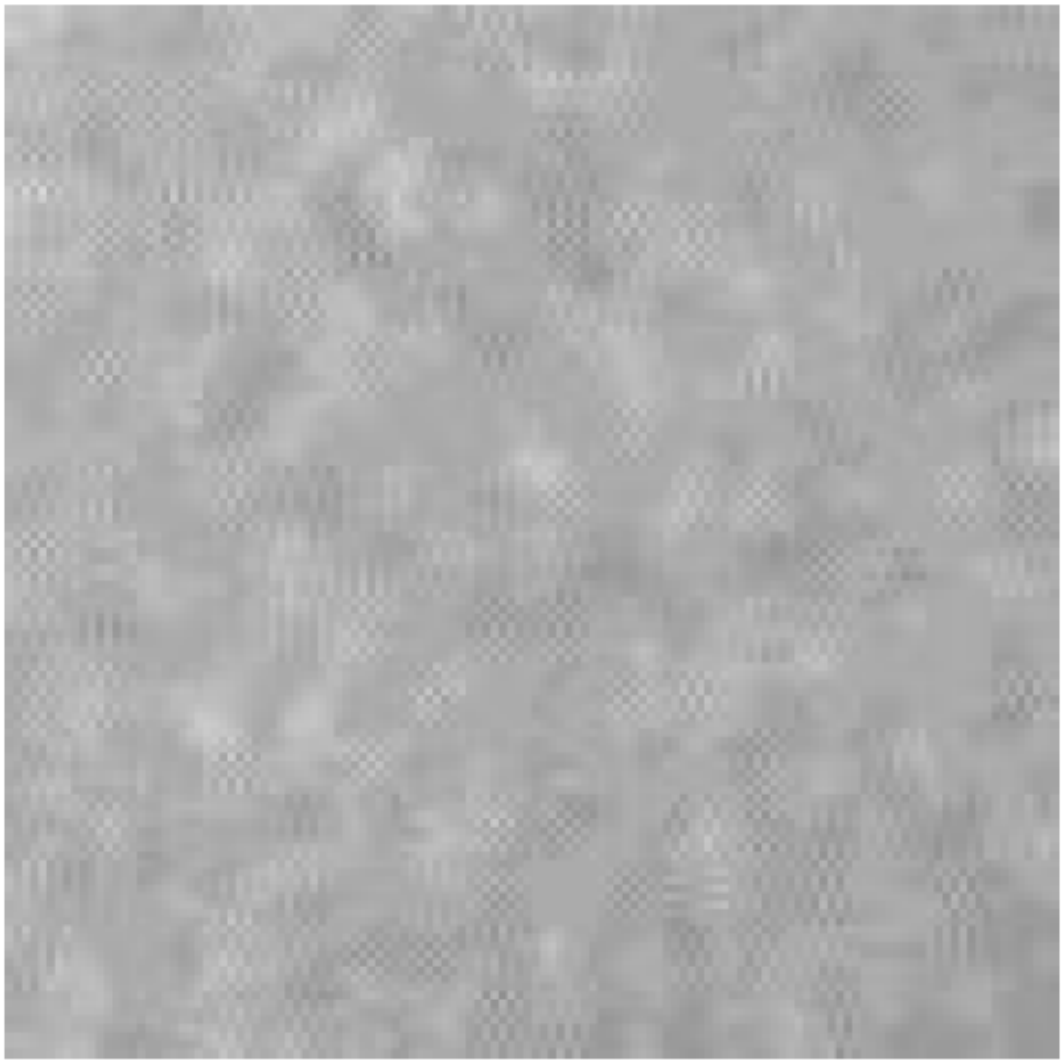


Coherent motion

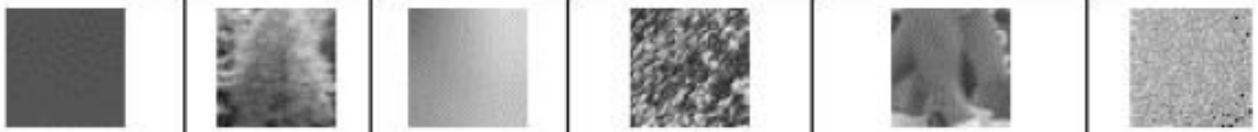
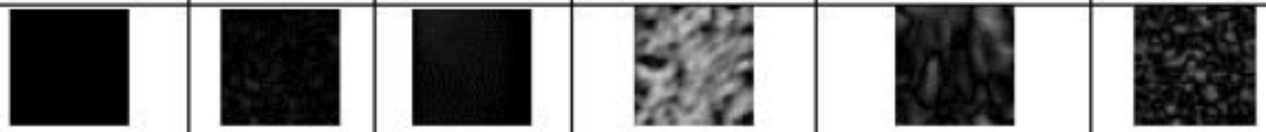
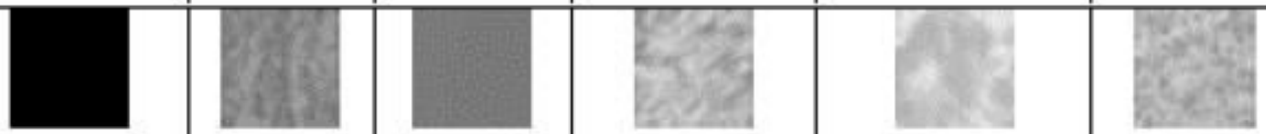
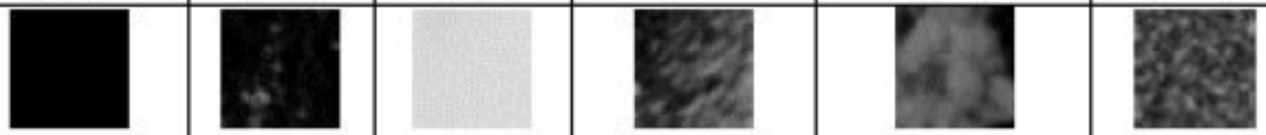


Incoherent motion

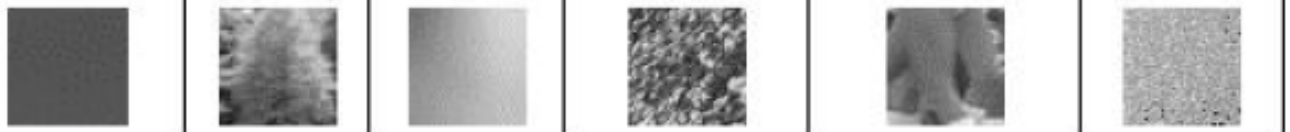
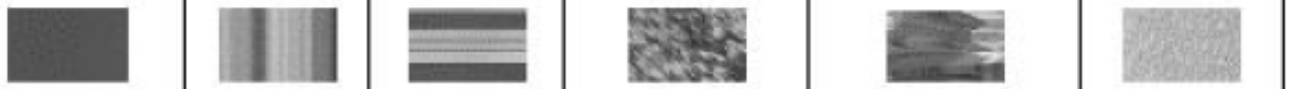
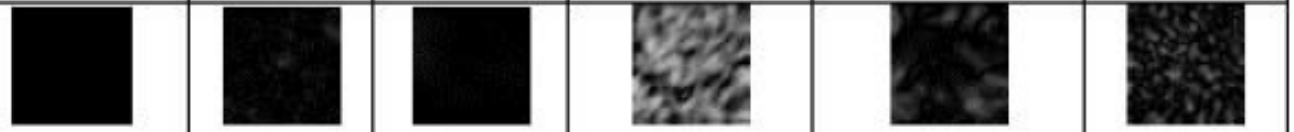
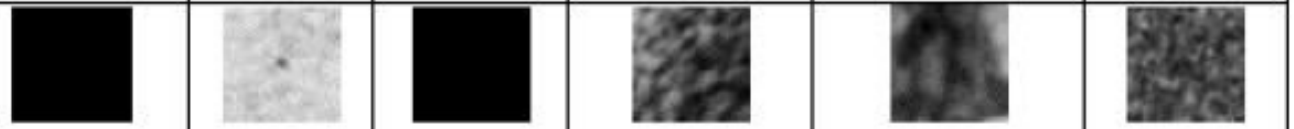
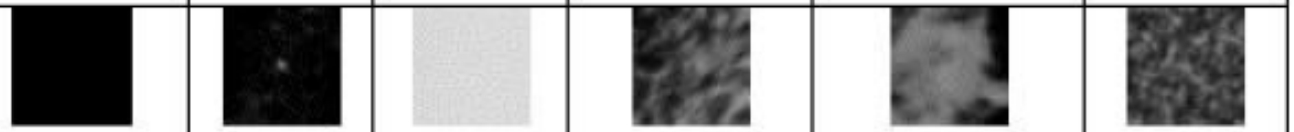
Scintillation



Example: Results x-t

	Unstruc-tured	Static	Flicker	Coherent Motion	Incoherent Motion	Scintil-lation
						
						
$\lvert R-L \rvert$	 0.00	0.00	0.00	0.37	0.05	0.02
$R+L$	 0.01	0.40	0.36	0.53	0.58	0.50
S_x	 0.00	0.55	0.00	0.21	0.17	0.25
F_x	 0.00	0.04	0.63	0.26	0.25	0.23

Example: Results y-t

	Unstruc-tured	Static	Flicker	Coherent Motion	Incoherent Motion	Scintil-lation
						
						
$\lvert U - D \rvert$	 0.00	0.00	0.00	0.34	0.02	0.02
$U + D$	 0.01	0.38	0.36	0.52	0.45	0.50
S_y	 0.00	0.59	0.00	0.19	0.24	0.28
F_y	 0.00	0.03	0.64	0.29	0.29	0.21

Outline

- Introduction
- Intuition
- Orientation in visual space-time
- A representation for spatiotemporal patterns
- Spatiotemporal boundaries
- Application to detection and tracking

Boundaries

- Zero-crossings in the double-opponent motion operator output indicate coherent motion boundaries.

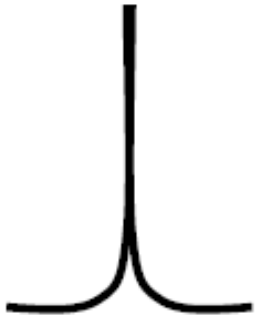
$$R - L$$

Boundaries

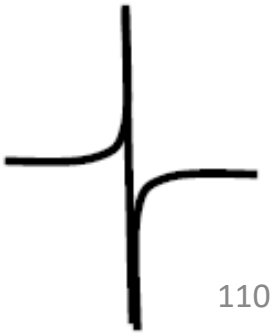
- Zero-crossings in the double-opponent motion operator output indicate coherent motion boundaries.
- Slope magnitude taken as strength of boundary signal.

$R - L$

$D_x(R - L)$



$D_{xx}(R - L)$

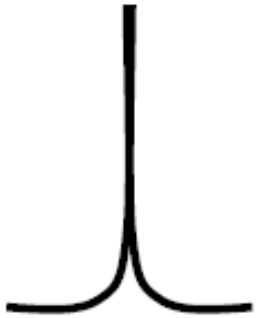


Boundaries

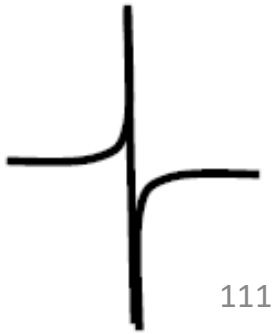
- Zero-crossings in the double-opponent motion operator output indicate coherent motion boundaries.
- Slope magnitude taken as strength of boundary signal.
- Sum signals from ***x-t*** and ***y-t*** dimensions.

$R - L$

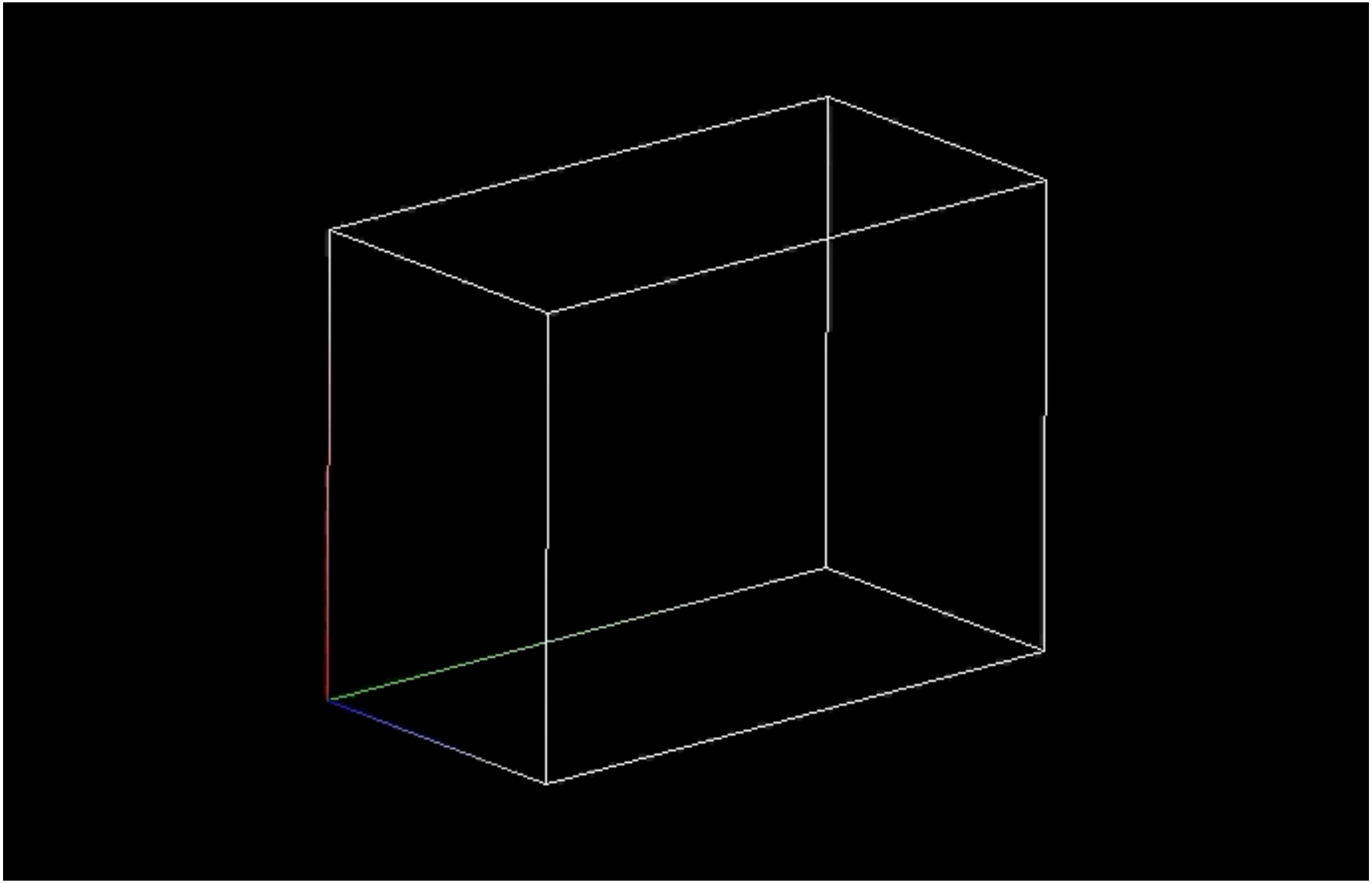
$D_x(R - L)$



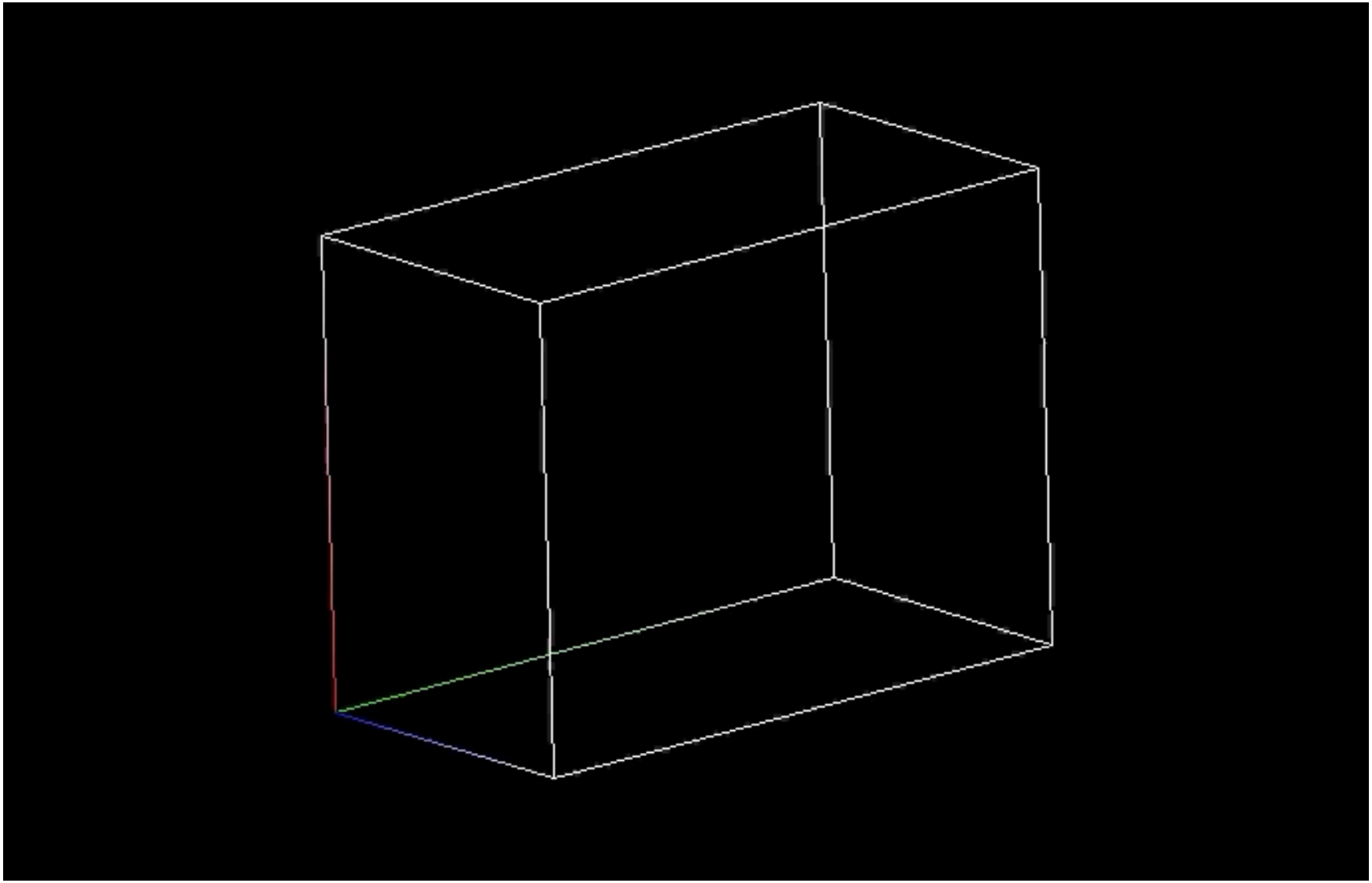
$D_{xx}(R - L)$



Pyramids



Pyramids



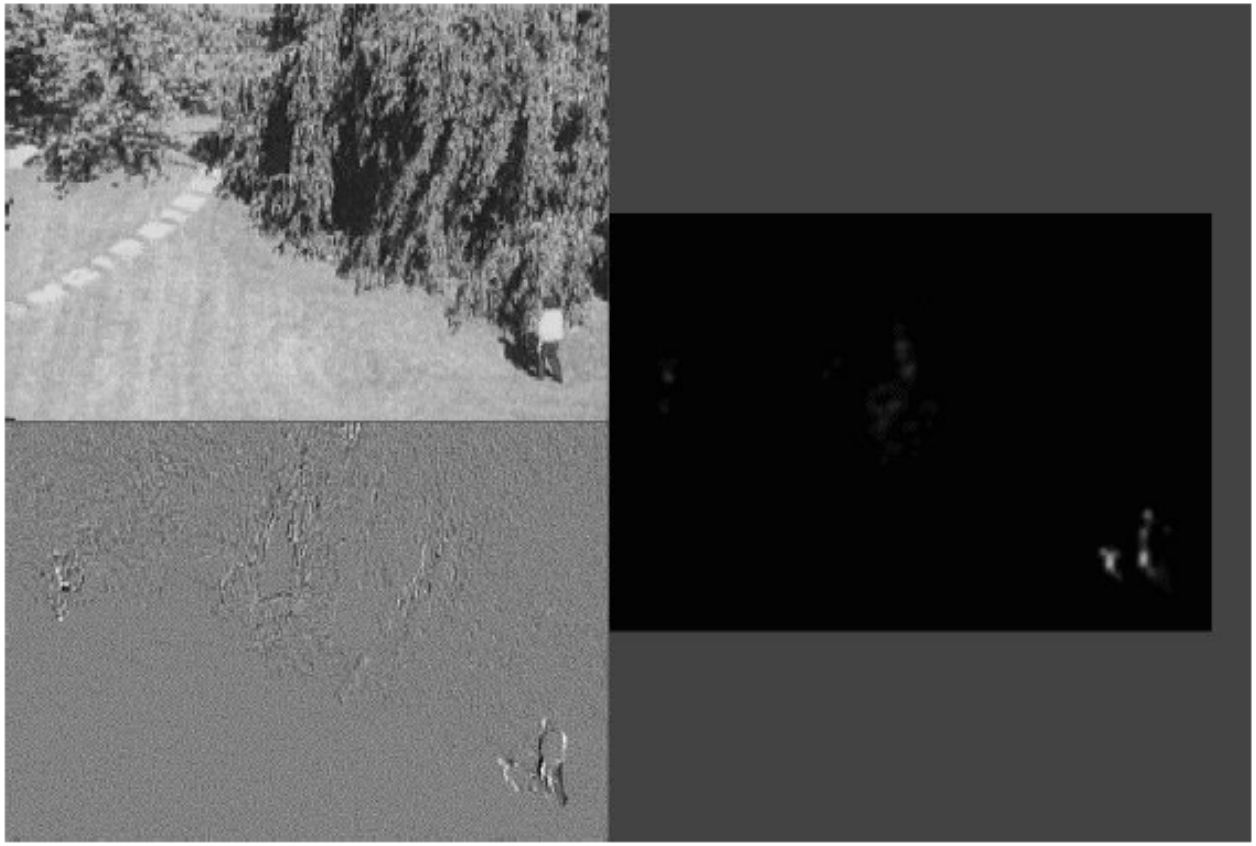
Outline

- Introduction
- Intuition
- Orientation in visual space-time
- A representation for spatiotemporal patterns
- Spatiotemporal boundaries
- Application to detection and tracking

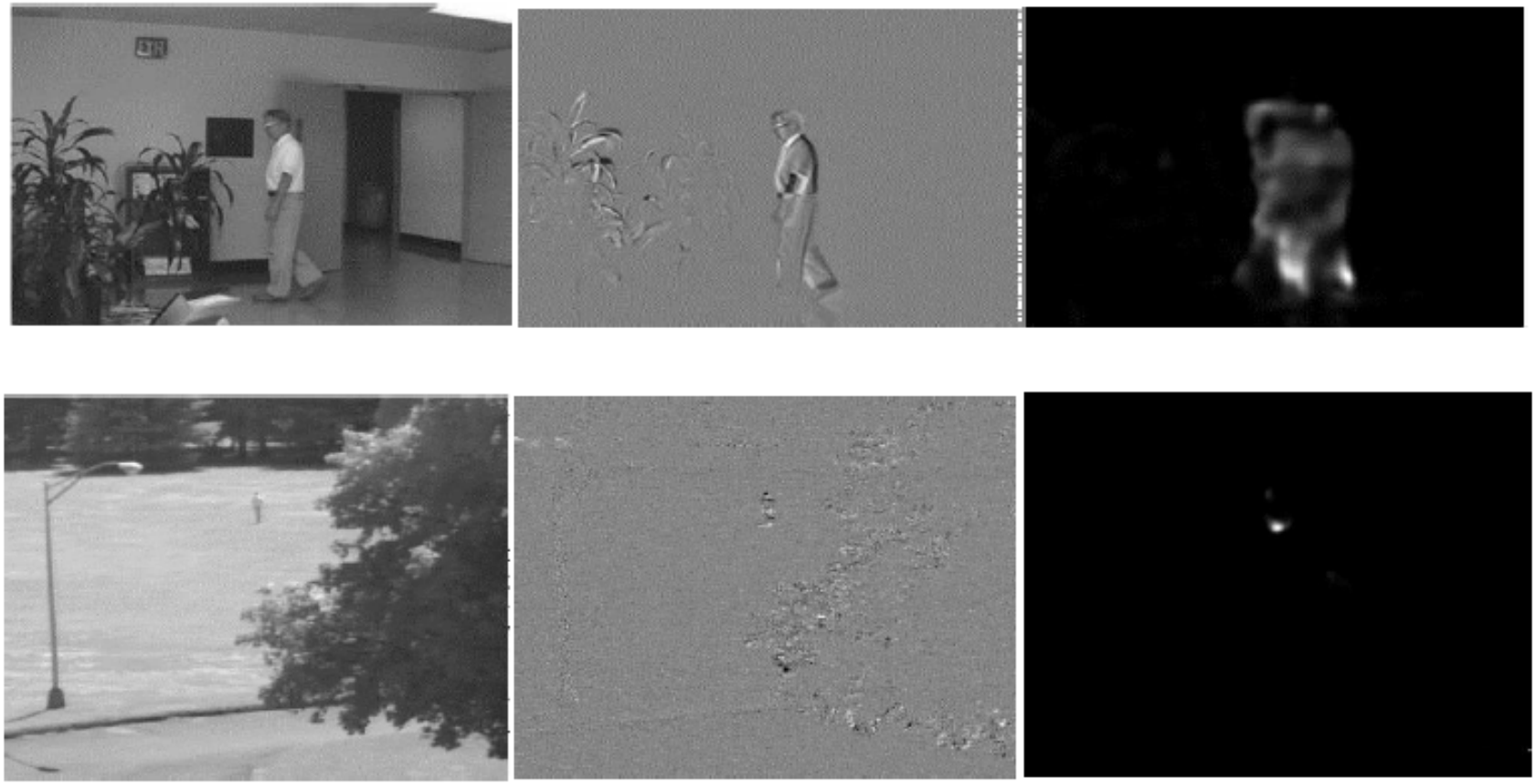
Detection and tracking

- Vision-based detection and tracking can be led astray by background clutter.
- Typical targets of interest exhibit motion that is locally coherent in space-time.
- Make use of proposed spatiotemporal representation of coherent motion as a cue to distinguish salient targets.

Detection and tracking



Detection and tracking



Summary

- Introduction
- Intuition
- Orientation in visual space-time
- A representation for spatiotemporal patterns
- Spatiotemporal boundaries
- Application to detection and tracking