Map-Seeking Circuit (MSC)

A Computational Mechanism for Recognition under Transformation with Algorithmic and Analog Implementations

David W. Arathorn

General Intelligence Corp

and

Montana State University Dept of Cell Biology and Neuroscience

General Intelligence

Original objective: computationally efficient, robust solution to the “correspondence problem” in vision.

Original objective: computationally efficient, robust solution to the “correspondence problem” in vision

Assumption: existence proof → *neuronally plausible operations and dataflow must constrain mathematical solution*.

Original objective: computationally efficient, robust solution to the “correspondence problem” in vision

Assumption: existence proof → *neuronally plausible operations and dataflow must constrain mathematical solution*.

Result: a general method for solution of a class of inverse problems that involve **transformation discovery**.

Original objective: computationally efficient, robust solution to the “correspondence problem” in vision

Assumption: existence proof → *neuronally plausible operations and dataflow must constrain mathematical solution*.

Result: a general method for solution of a class of inverse problems that involve **transformation discovery**.

By-product result: an analog circuit with neuron-like elements for implementing that method of solution.

Original objective: computationally efficient, robust solution to the "correspondence problem" in vision

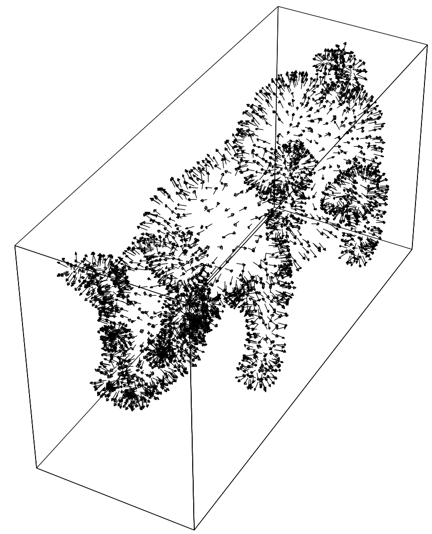
Assumption: existence proof → *neuronally plausible operations and dataflow must constrain mathematical solution*.

Result: a general method for solution of a class of inverse problems that involve **transformation discovery**.

By-product result: an analog circuit with neuron-like elements for implementing that method of solution. Is it brain-like?

Applicable to vision, inverse kinematics, route finding, puzzle solving(Rubik's cube), <u>cone-targeted stimulus delivery</u>, other…

The “correspondence problem” in vision



Separating objects of interest from other objects or background, *segmentation*, is an emergent consequence of transformation discovery.

Transformation Discovery

Versus

Transformation "Invariance"

Discovery = Aware of Transformations, reports:

Object identity (class) and

Parameters of Image Formation, and

Object Transformations (e.g. pose, morphology)

Invariance = Blind to Transformations, reports:

Object identity (class) only

Successful “invariance” is blind to the difference in pose, position, and scale (and variations in morphology).

Psychophysics (and common experience) is the falsifying experiment.

Correspondence is a measure of how well an input pattern **r** matches a memory template **w** after **r** has been subjected to a sequence of transformations, t_{j1}, t_{j2}… t_{jL}.

$$c(\mathbf{j}) = \left\langle t_{jL}^{L} \circ \cdots \circ t_{j2}^{2} \circ t_{j1}^{1}(\mathbf{r}), \mathbf{w} \right\rangle$$

or more compactly

$$c(\mathbf{j}) = \left\langle \mathop{\circ}_{l=1}^{L} t_{jl}^{l}(\mathbf{r}), \mathbf{w} \right\rangle$$

MSC solves

$$\boxed{\mathbf{x} = \arg\max c(\mathbf{j}) = \left\langle \mathop{\circ}_{i=1}^{L} t_{j_i}^{i}(\mathbf{r}), \mathbf{w}_{j_{L+1}}^{L+1} \in \mathbf{W} \right\rangle}$$

i.e. optimizes selection of transformation sequence and memory template

Accommodating transformations

Transformations due to

Image formation: location, distance, orientation, occlusion

Object configuration: morphological variation, articulation

SOURCES OF BIO-INSPIRED COMPUTATIONAL HARDWARE

Direct engineering from biological prototype

Desirable but unlikely (see C. Koch 2013 presentation)

Reverse-engineering deduced from multiple available constraints

1. *Functionality, psychophysics*

2. *Neuroanatomy, neurophysiology*

3. Plausible neuronal computational repertoire

1. Functionality and psychophysics

Not just the obvious… e.g. *human vision capable of classification at 4-8 cycles on object long axis (Army NightVisionLab).*

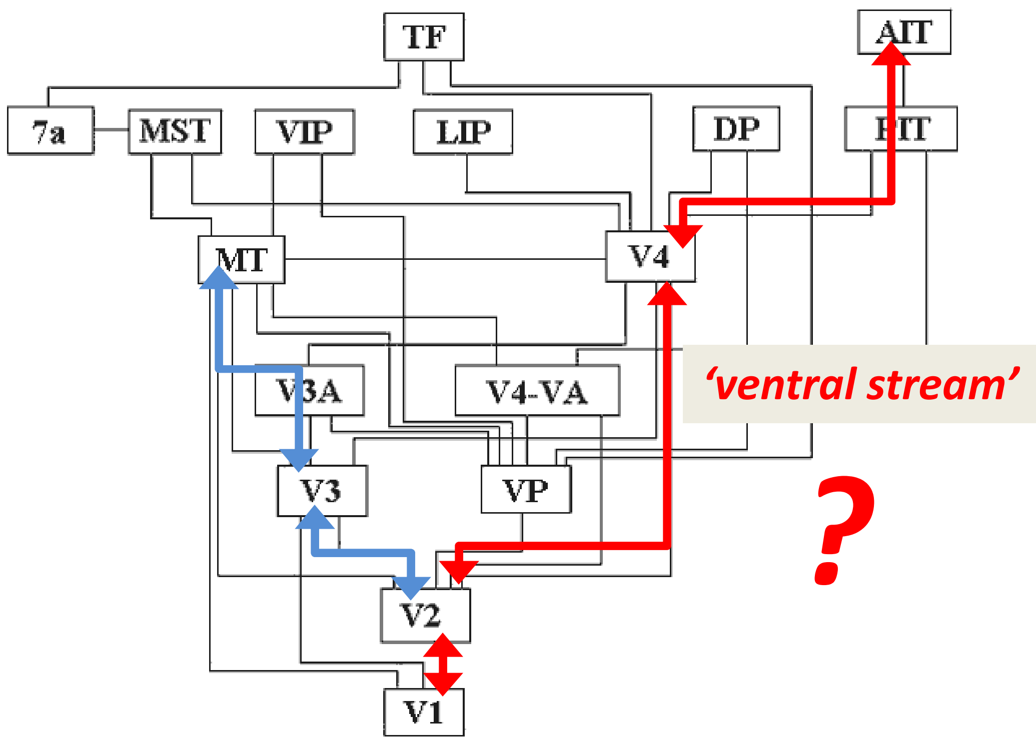
Target size is the dominant factor in filter-type object detector performance…

For lowest extent decile, recognition rates between 0.0 and 0.1

From ***Diagnosing Error in Object Detectors,***
Derek Hoiem, Yodsawalai Chodpathumwan, and
Qieyun Dai, ECCV,2013

2. Neuroanatomy, neurophysiology

Bi-directional datapaths (more top-down than bottom-up).
Inter-area connectivity patterns [Angelucci et al].
Flow through multiple stages of area-specific functionality.



3. Neuron computational repertoire

Plausible neuronal computational operations

Combination: some monotonic function of sum of signals $\beta\left(\sum_i a_i\right)$

Gating: an analog scaling of one signal by another $g_j \cdot a_i$ [B. Mel]

Mapping: reordering of spatial arrangement of signals $F\colon \mathbf{x} \to \mathbf{x}'$

Inhibition: attenuation of one signal by another $g_j^{-1} \cdot a_i$ or $a_i - b_j$

…**gating** and **combination** ⇨ weighted **superpositions**

…**gating** and **combination** ⇨ **matching** operation (e.g. dot product)

…**mapping** ⇨ geometric **transforms**

…**inhibition** ⇨ **competition**

Crux of Correspondence Problem: Combinatorial Explosion…

Pixel positions in 512 x 512 image = 400^2

Scalings = 30 (assuming rough knowledge of distance)

Rotations in viewing plane = 30 (assuming normal target orientation)

Azimuth rotations = 36 (at 10 degree increments)

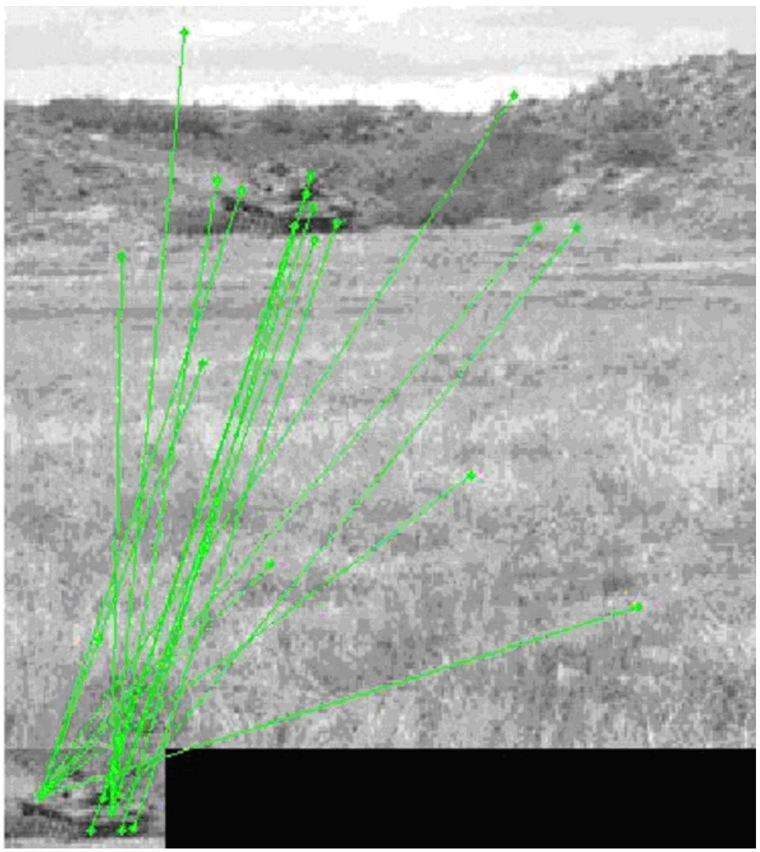
Elevation rotations = 10 (at 10 degree increments)

5.184 x 10^{10} possible distinct transformations

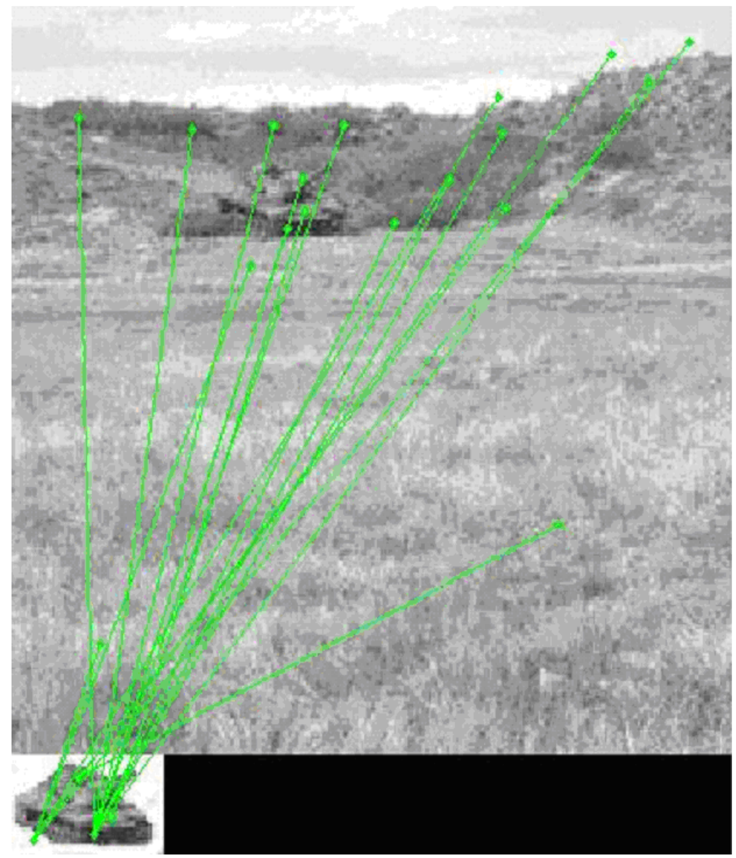
Usual MV Suspect: SIFT + RANSAC + Fundamental Matrix

Requires ***extreme similarity*** between template and image.
And not plausible for neuronal solution.

Object AND background match



Object only match



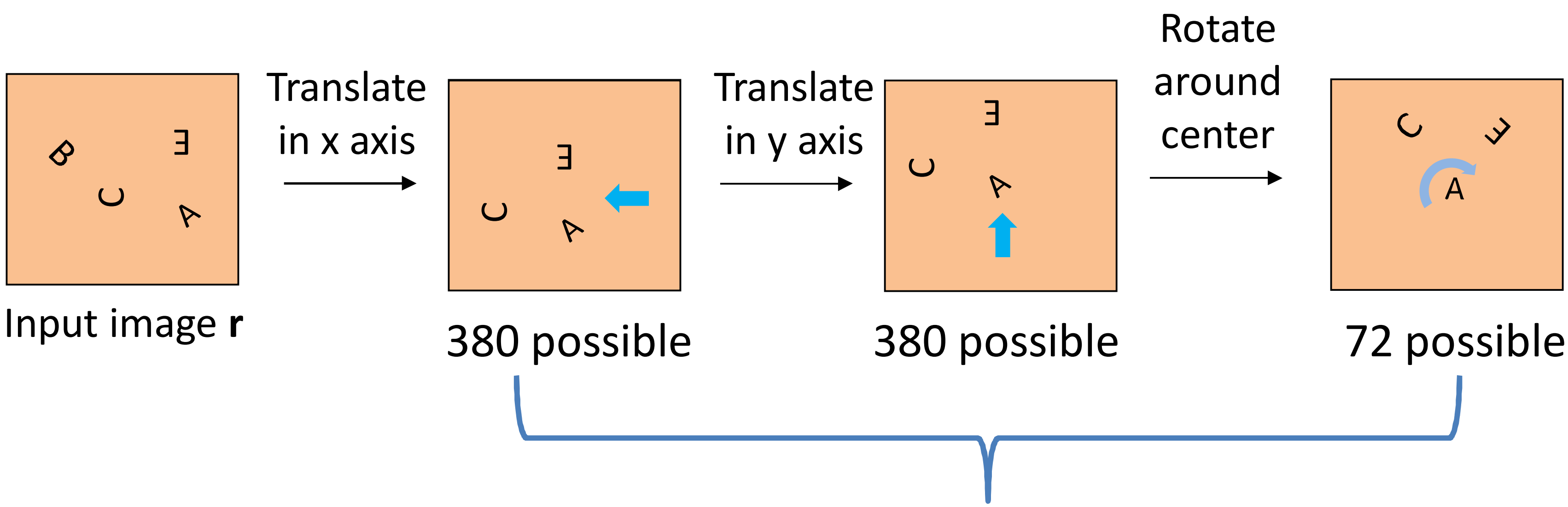
How do we efficiently recognize the rotated “A”?

Assume input image is 400 pixels by 400 pixels

Assume the “**A**” can be rotated in 5 degree increments

By brute force, **over 10 million combinations** of translation and rotation.

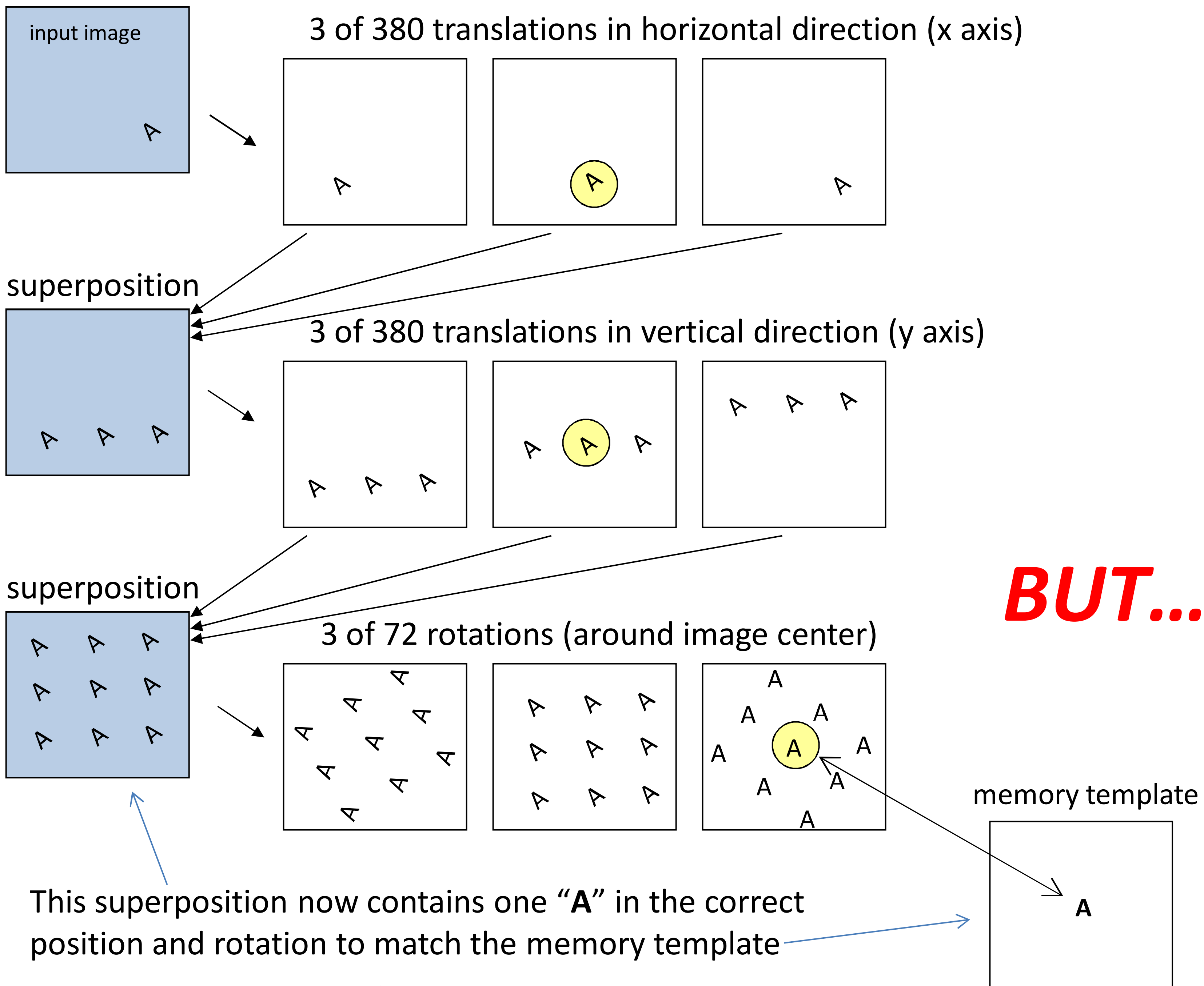
However, if we *decompose* all possible combinations into three *stages* (or *layers*)…



sum of the possible transformations in each stage
instead of the
product of the number in each stage…

Or **832** **versus > 10 million** ***IF…***

Apply transformations to superpositions of transforms…


input image
3 of 380 translations in horizontal direction (x axis)
superposition
3 of 380 translations in vertical direction (y axis)
superposition
3 of 72 rotations (around image center)
BUT…
memory template
This superposition now contains one “**A**” in the correct position and rotation to match the memory template

Bi-directional transformation

and

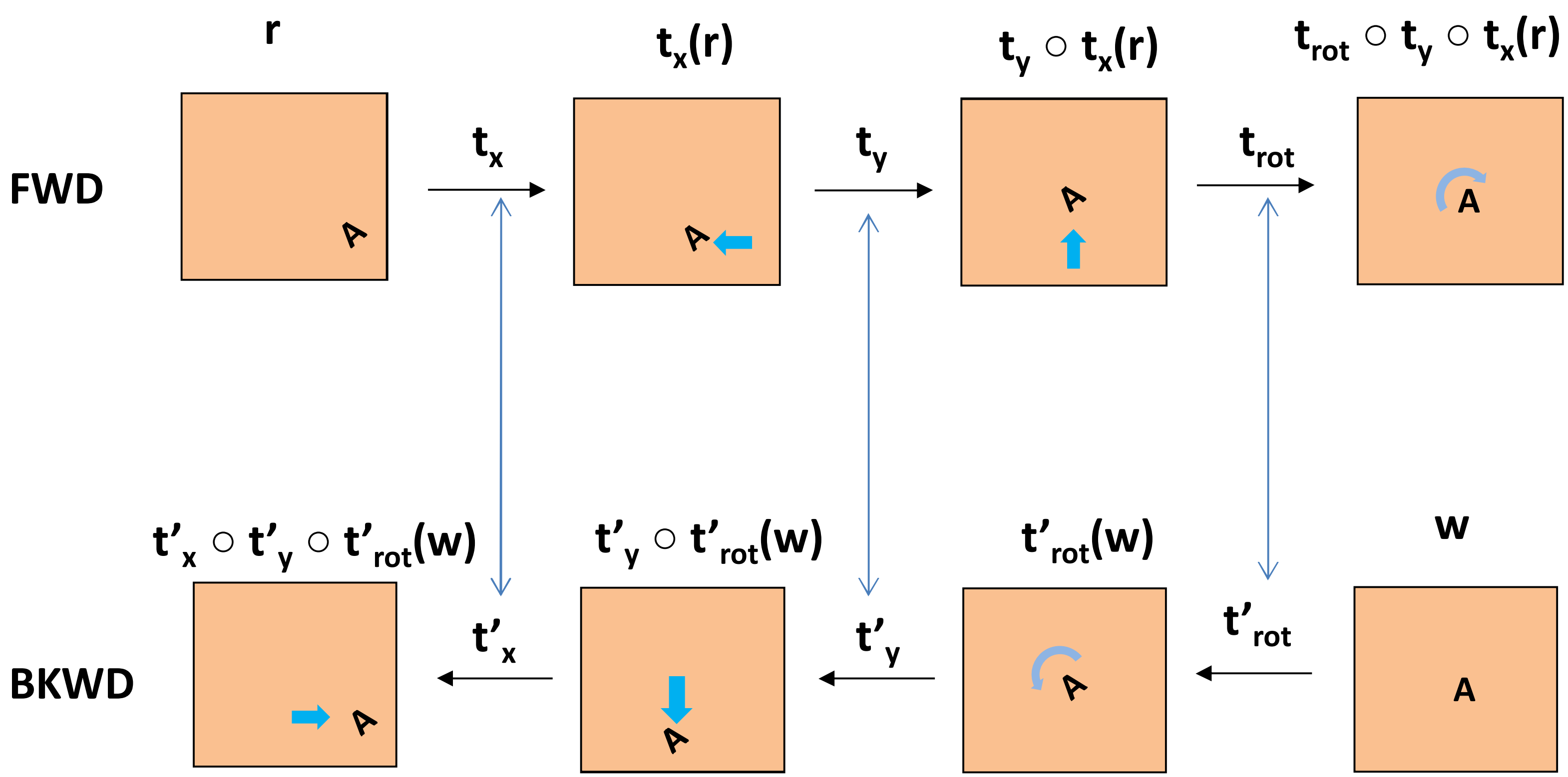
competition

and

ordering property of superpositions

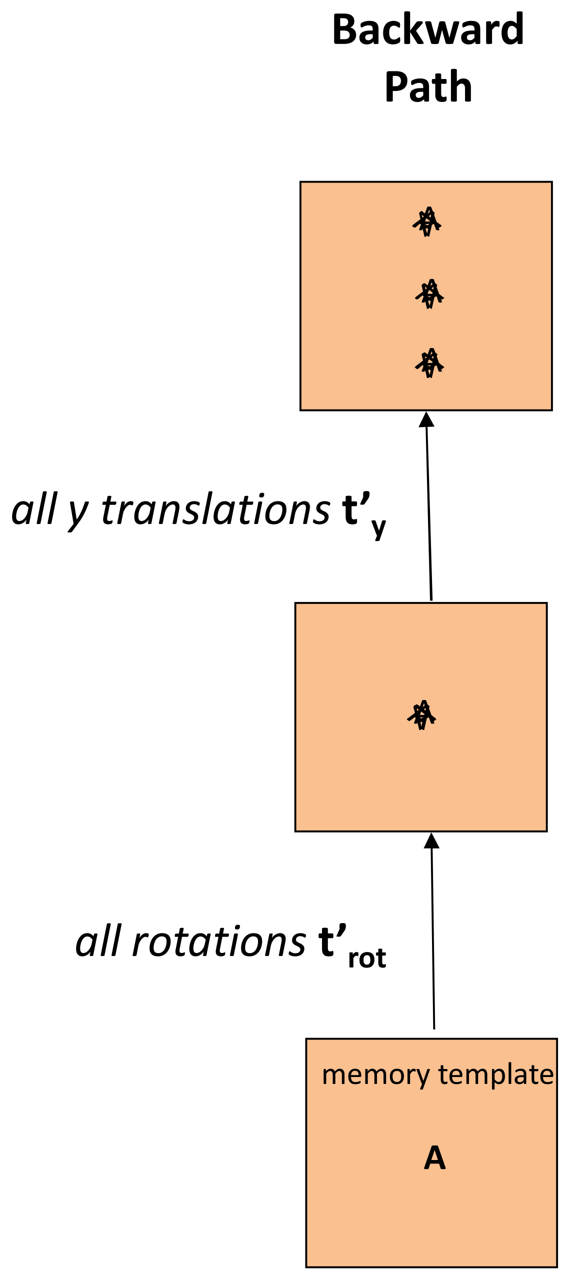
For each transformation step from **r** to **w…**
an inverse transformation step from **w** to **r.**

(The inverse of transformation **t** is notated as **t’)**



Notice that each vertical pair of images matches exactly.

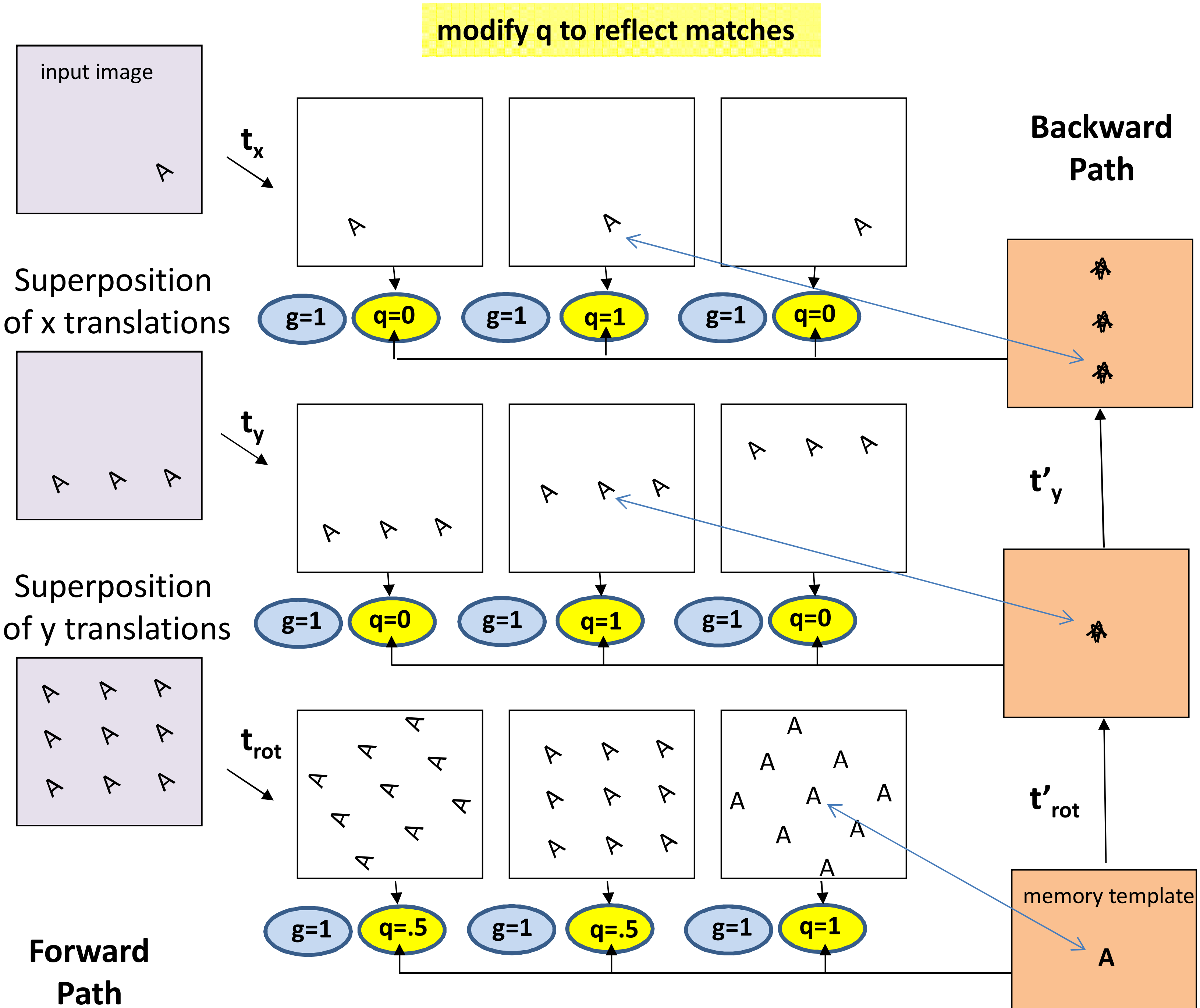
Inverse transformations applied to the memory template (**w**) form superpositions on the **backward path**.



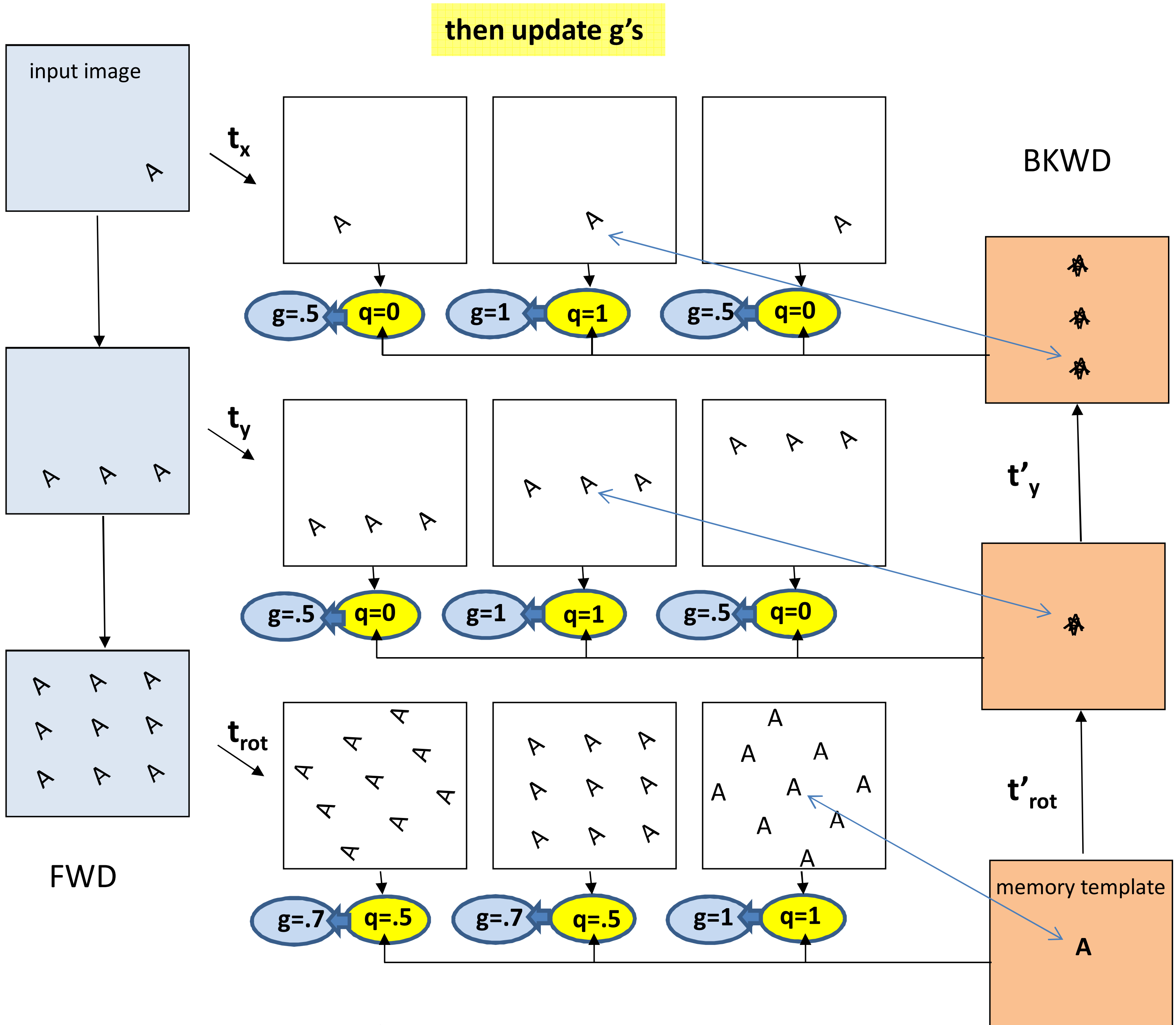
Gating and **matching operations** steer the selection of the correct transformation

q = degree of matching or correlation

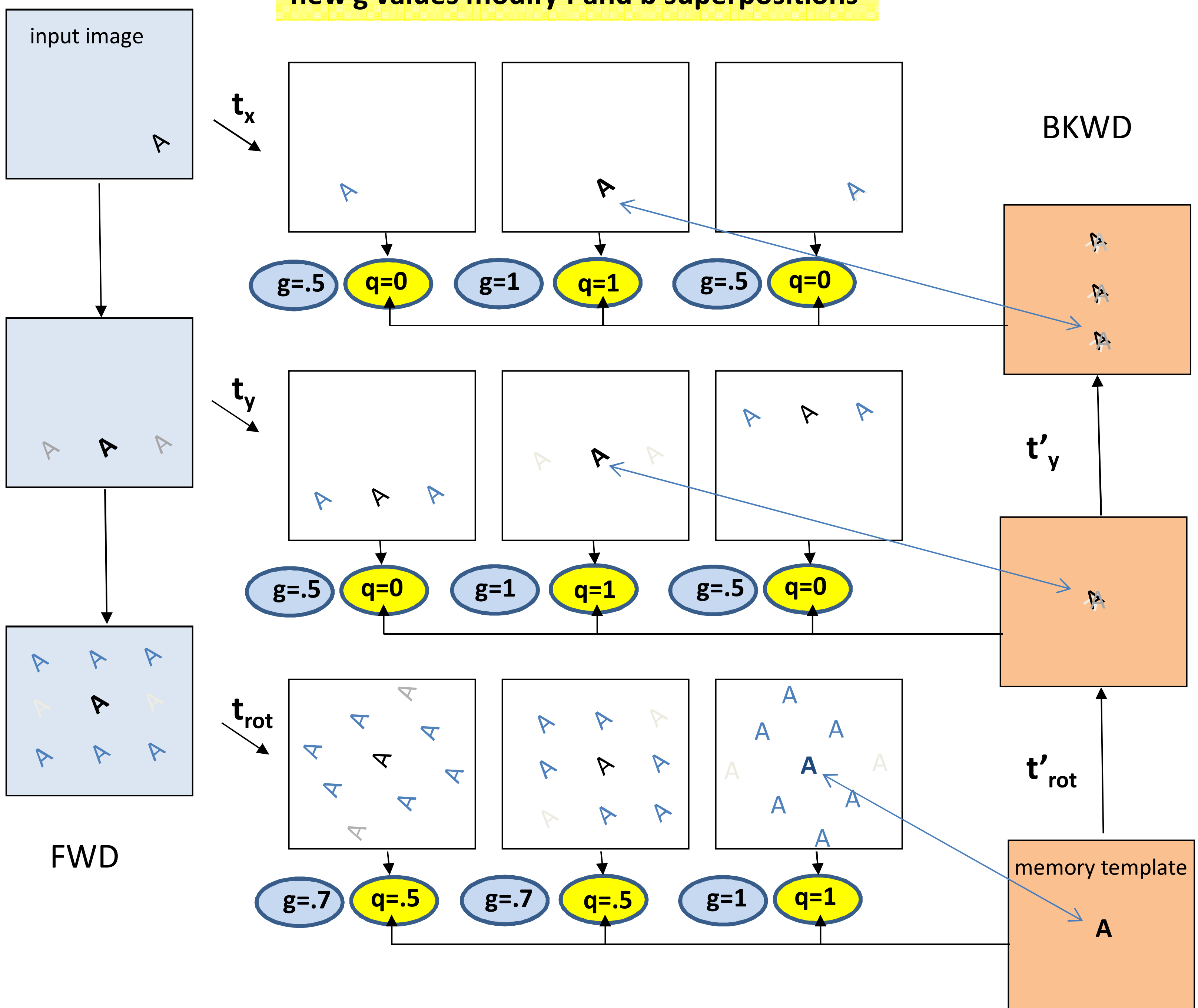
g = gating coefficient

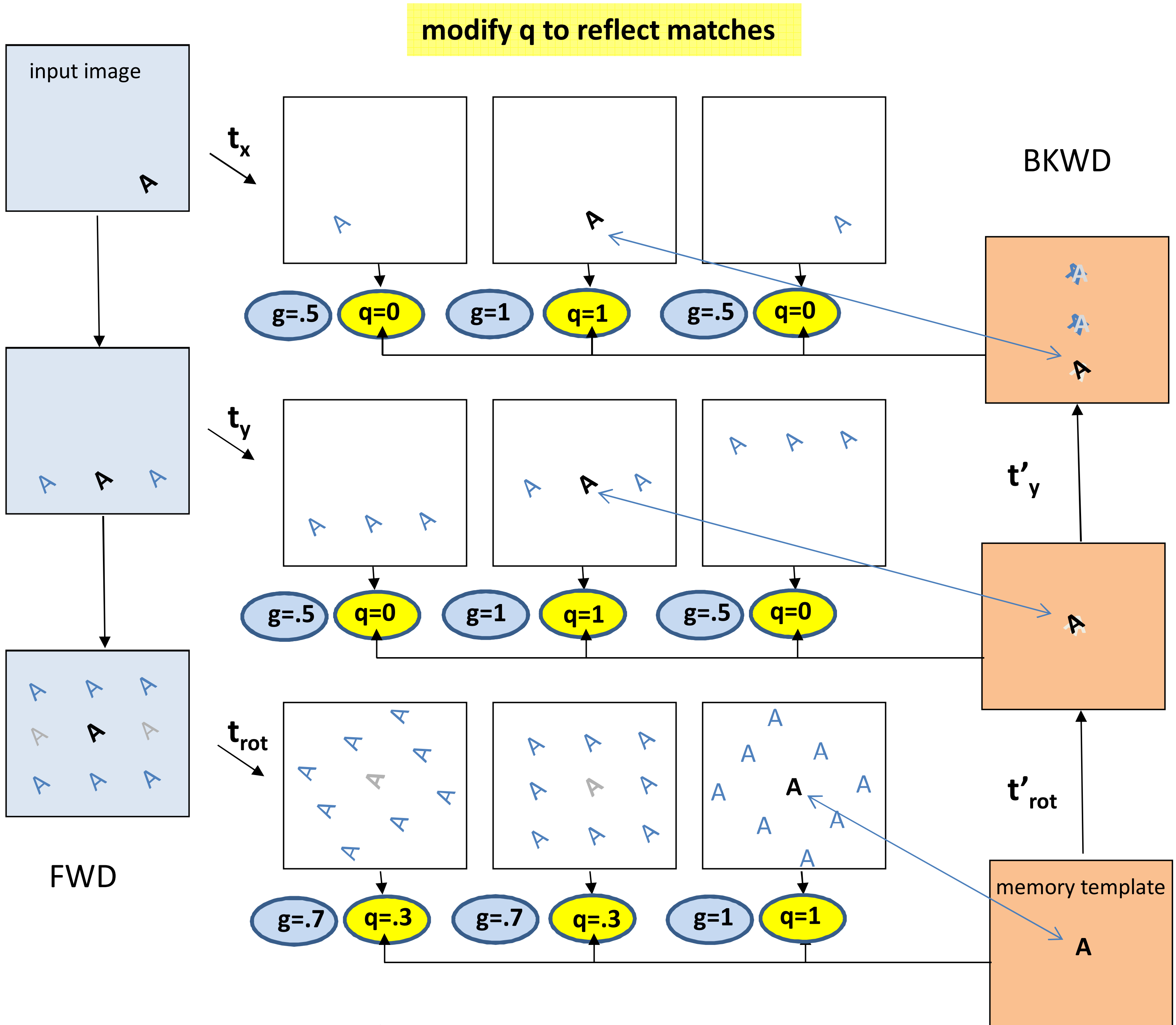


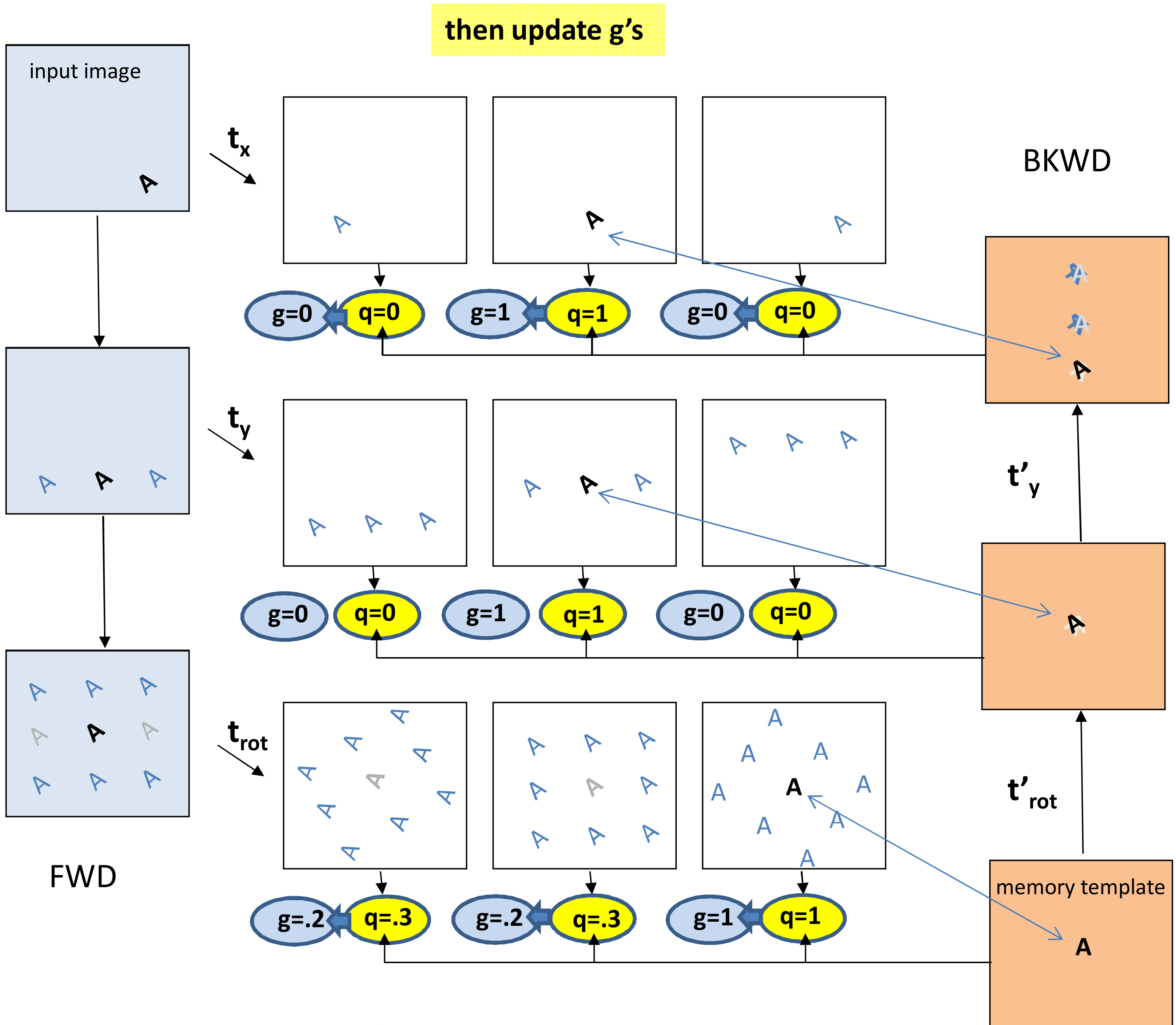
modify q to reflect matches
input image
t_x
Backward Path
Superposition of x translations
g=1
q=0
g=1
q=1
g=1
q=0
t_y
t'_y
Superposition of y translations
g=1
q=0
g=1
q=1
g=1
q=0
t_{rot}
t'_{rot}
g=1
q=.5
g=1
q=.5
g=1
q=1
memory template
A
Forward Path

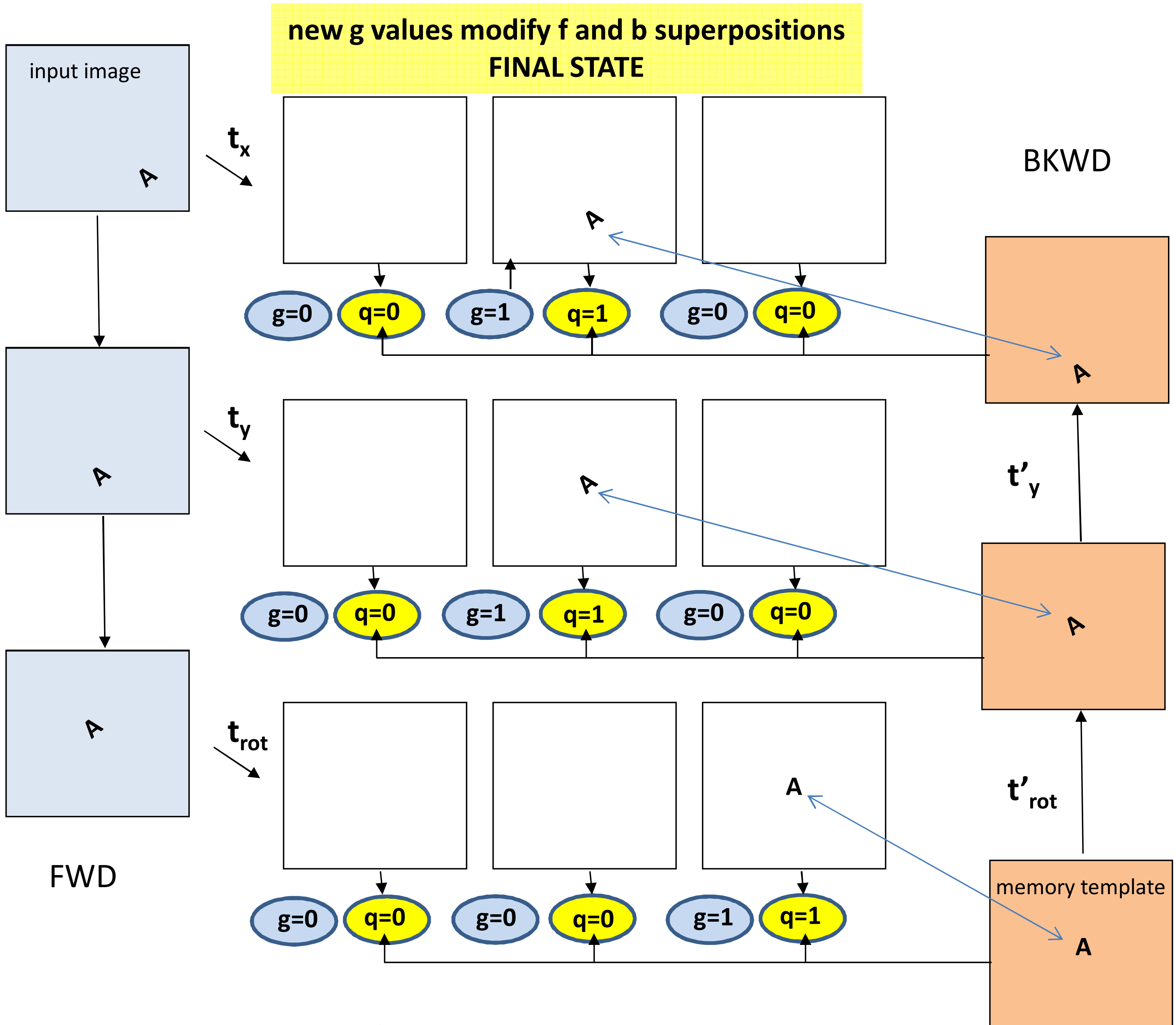

then update g's
input image
t_x
t_y
t_{rot}
BKWD
FWD
g=.5
q=0
g=1
q=1
g=.5
q=0
g=.5
q=0
g=1
q=1
g=.5
q=0
g=.7
q=.5
g=.7
q=.5
g=1
q=1
t'_y
t'_{rot}
memory template
A

new g values modify f and b superpositions

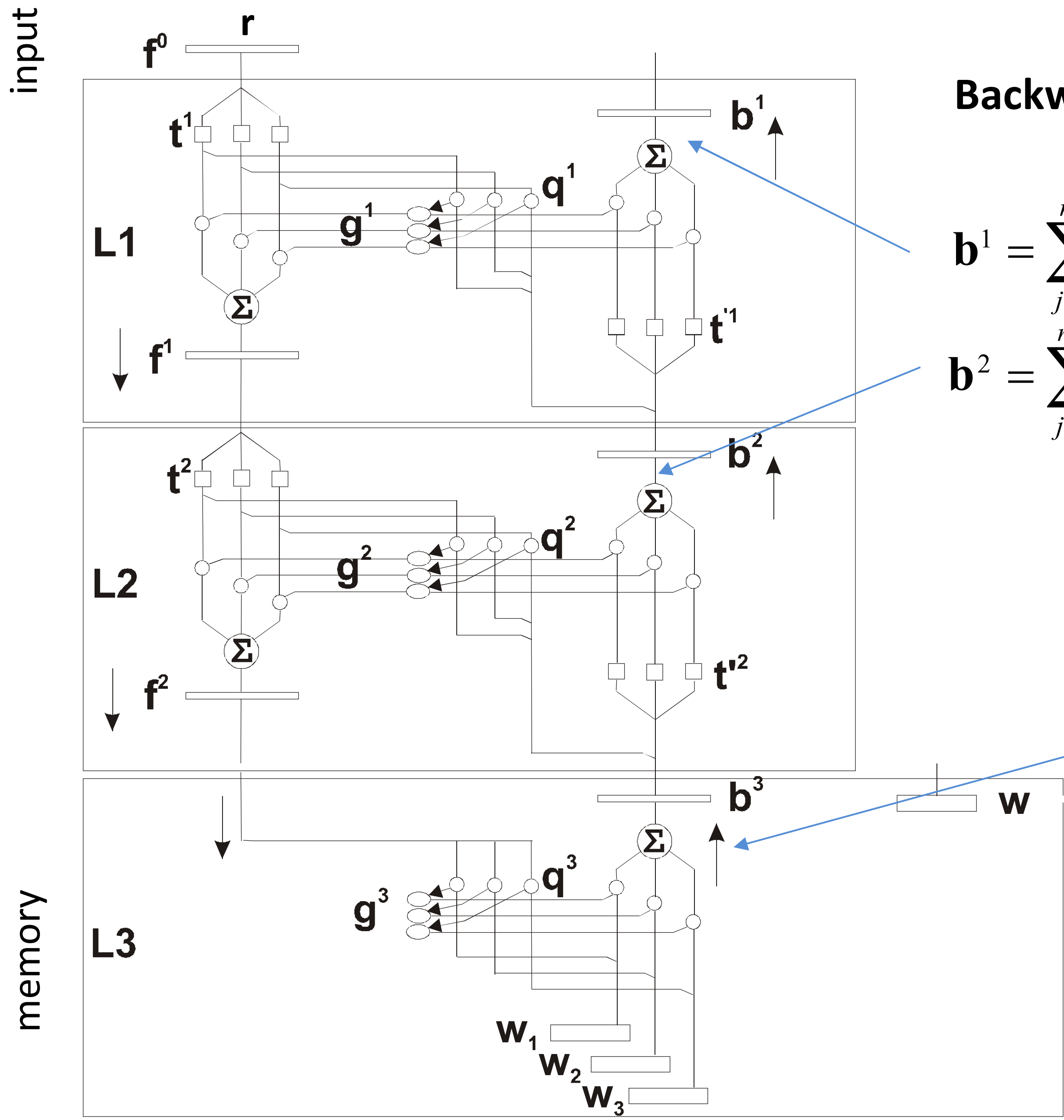



modify q to reflect matches
input image
tx
ty
trot
FWD
BKWD
t'y
t'rot
memory template
g=.5
q=0
g=1
q=1
g=.5
q=0
g=.5
q=0
g=1
q=1
g=.5
q=0
g=.7
q=.3
g=.7
q=.3
g=1
q=1


then update g's
input image
FWD
BKWD
t_x
t_y
t_{rot}
t'_y
t'_{rot}
memory template
g=0
q=0
g=1
q=1
g=0
q=0
g=0
q=0
g=1
q=1
g=0
q=0
g=.2
q=.3
g=.2
q=.3
g=1
q=1


new g values modify f and b superpositions
FINAL STATE
input image
A
t_x
A
g=0
q=0
g=1
q=1
g=0
q=0
BKWD
A
A
t_y
A
g=0
q=0
g=1
q=1
g=0
q=0
t'_y
A
A
t_rot
A
g=0
q=0
g=0
q=0
g=1
q=1
t'_rot
memory template
A
FWD

Backward Path Computations



$$\mathbf{b}^1 = \sum_{j=1}^{n_1} g_j^1 \cdot t_j'^1\left(\mathbf{b}^2\right)$$

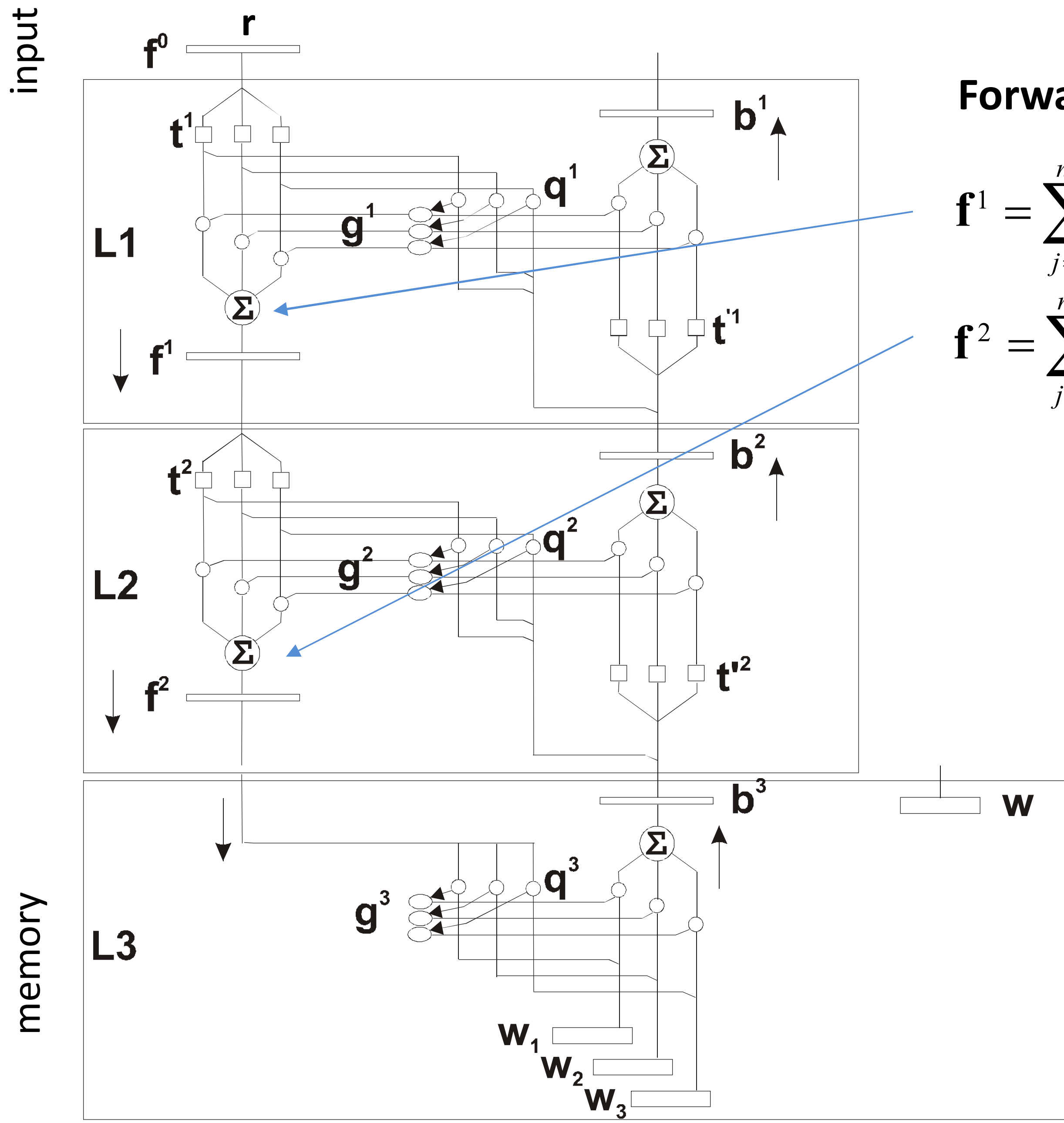
$$\mathbf{b}^2 = \sum_{j=1}^{n_2} g_j^2 \cdot t_j'^2\left(\mathbf{b}^3\right)$$

$$\mathbf{b}^3 = \sum_{k} g_k^3 \cdot \mathbf{w}_k$$

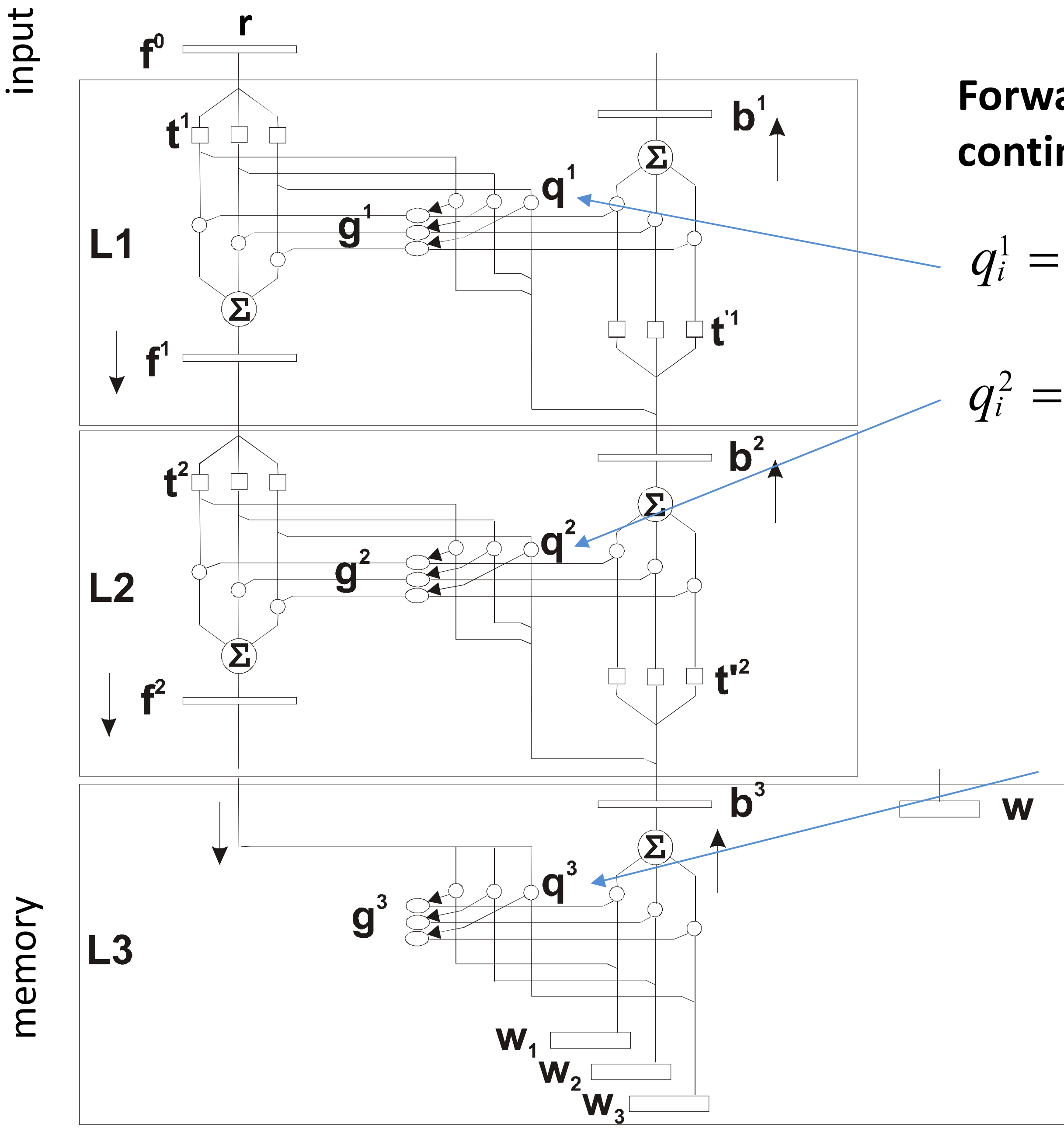
Forward Path Computations

$$\mathbf{f}^1 = \sum_{j=1}^{n1} g_j^1 \cdot t_j^1\left(\mathbf{f}^0\right)$$

$$\mathbf{f}^2 = \sum_{j=1}^{n2} g_j^2 \cdot t_j^2\left(\mathbf{f}^1\right)$$



Forward Path Computations, continued

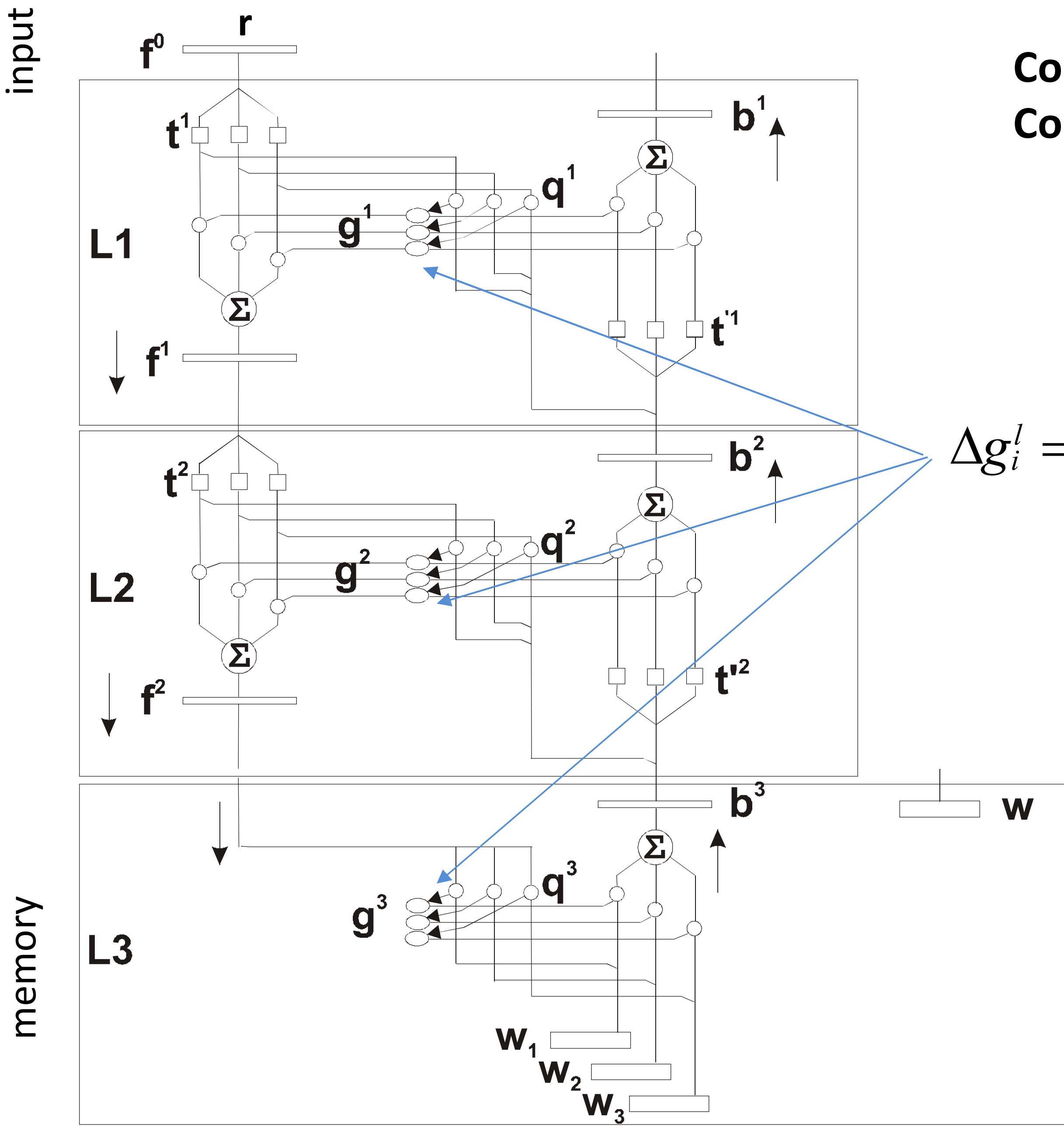


$$q_i^1 = \left\langle t_i^1\left(\mathbf{f}^0\right), \mathbf{b}^2 \right\rangle$$

$$q_i^2 = \left\langle t_i^2\left(\mathbf{f}^1\right), \mathbf{b}^3 \right\rangle$$

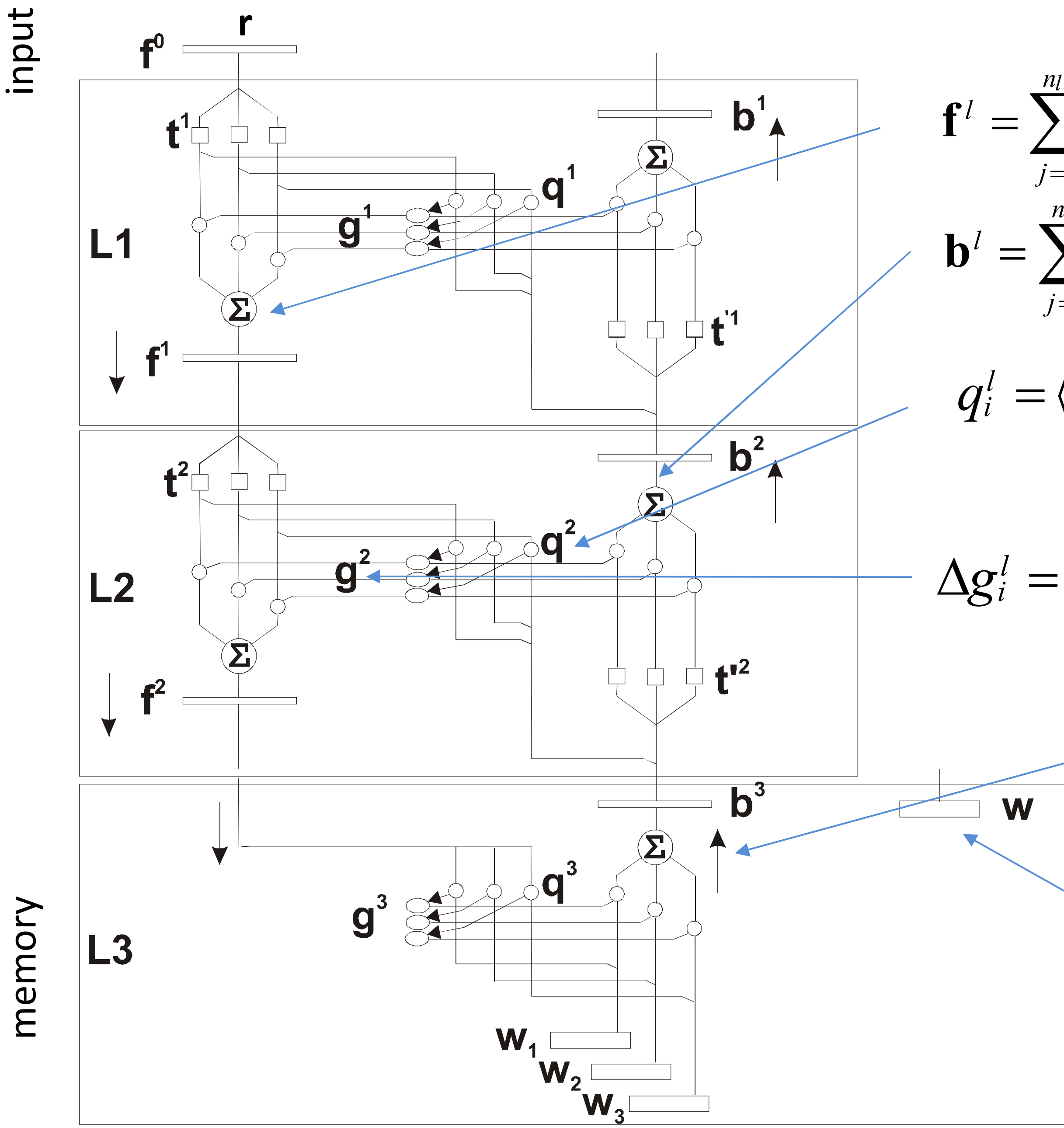
$$q_j^3 = \left\langle \mathbf{f}^2, \mathbf{w}_j \right\rangle$$

Competition and Gain Computations



$$\Delta g_i^l = -k_l \left(1 - \frac{q_i^l}{\max\left(\mathbf{q}^l\right)} \right)$$

$l = 1 \ldots 3$



$$\mathbf{f}^{l} = \sum_{j=1}^{n_l} g_j^l \cdot t_j^l\left(\mathbf{f}^{l-1}\right)$$

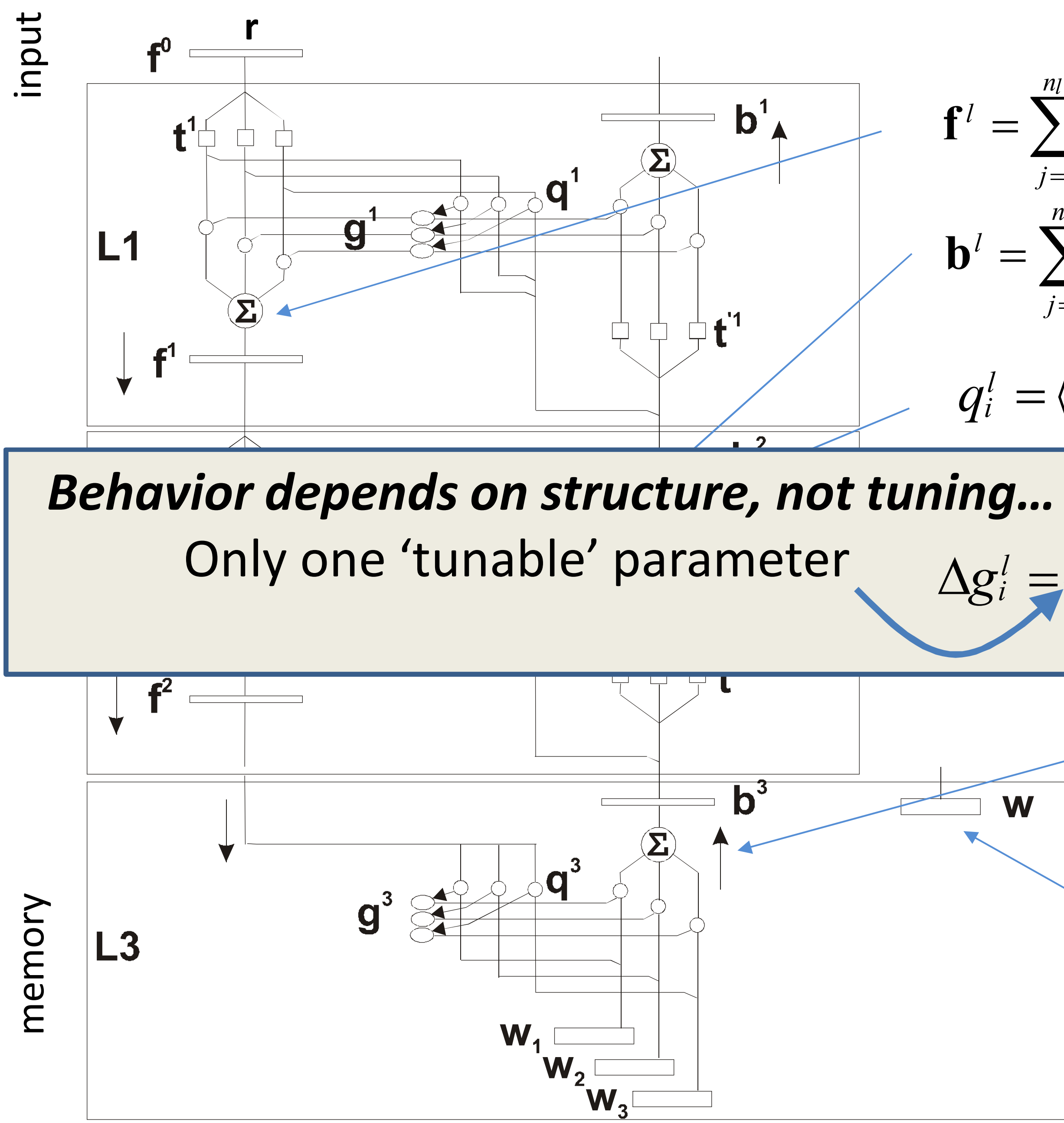
$$\mathbf{b}^{l} = \sum_{j=1}^{n_l} g_j^l \cdot t_j'^{l}\left(\mathbf{b}^{l+1}\right)$$

$$q_i^l = \left\langle t_i^l\left(\mathbf{f}^{l-1}\right), \mathbf{b}^{l+1} \right\rangle$$

$$\Delta g_i^l = -k_l\left(1 - \frac{q_i^l}{\max\left(\mathbf{q}^l\right)}\right)$$

$$\mathbf{b}^{3} = \sum_{k} g_k^3 \cdot \mathbf{w}_k$$

or $\mathbf{b}^3 = \mathbf{w}$



$$\mathbf{f}^{l} = \sum_{j=1}^{n_l} g_j^l \cdot t_j^l\left(\mathbf{f}^{l-1}\right)$$

$$\mathbf{b}^{l} = \sum_{j=1}^{n_l} g_j^l \cdot t_j'^l\left(\mathbf{b}^{l+1}\right)$$

$$q_i^l = \left\langle t_i^l\left(\mathbf{f}^{l-1}\right), \mathbf{b}^{l+1} \right\rangle$$

Behavior depends on structure, not tuning…

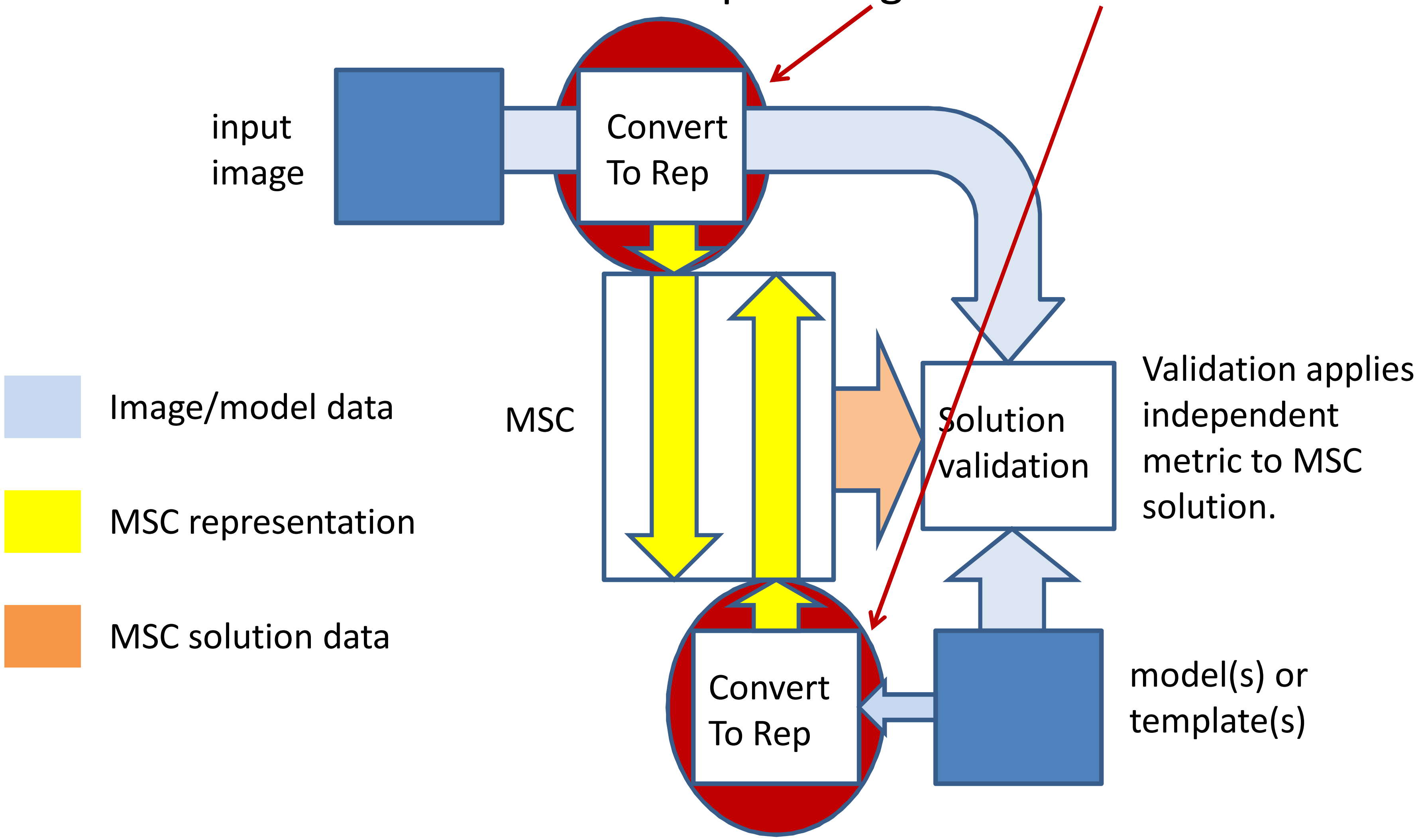
Only one 'tunable' parameter

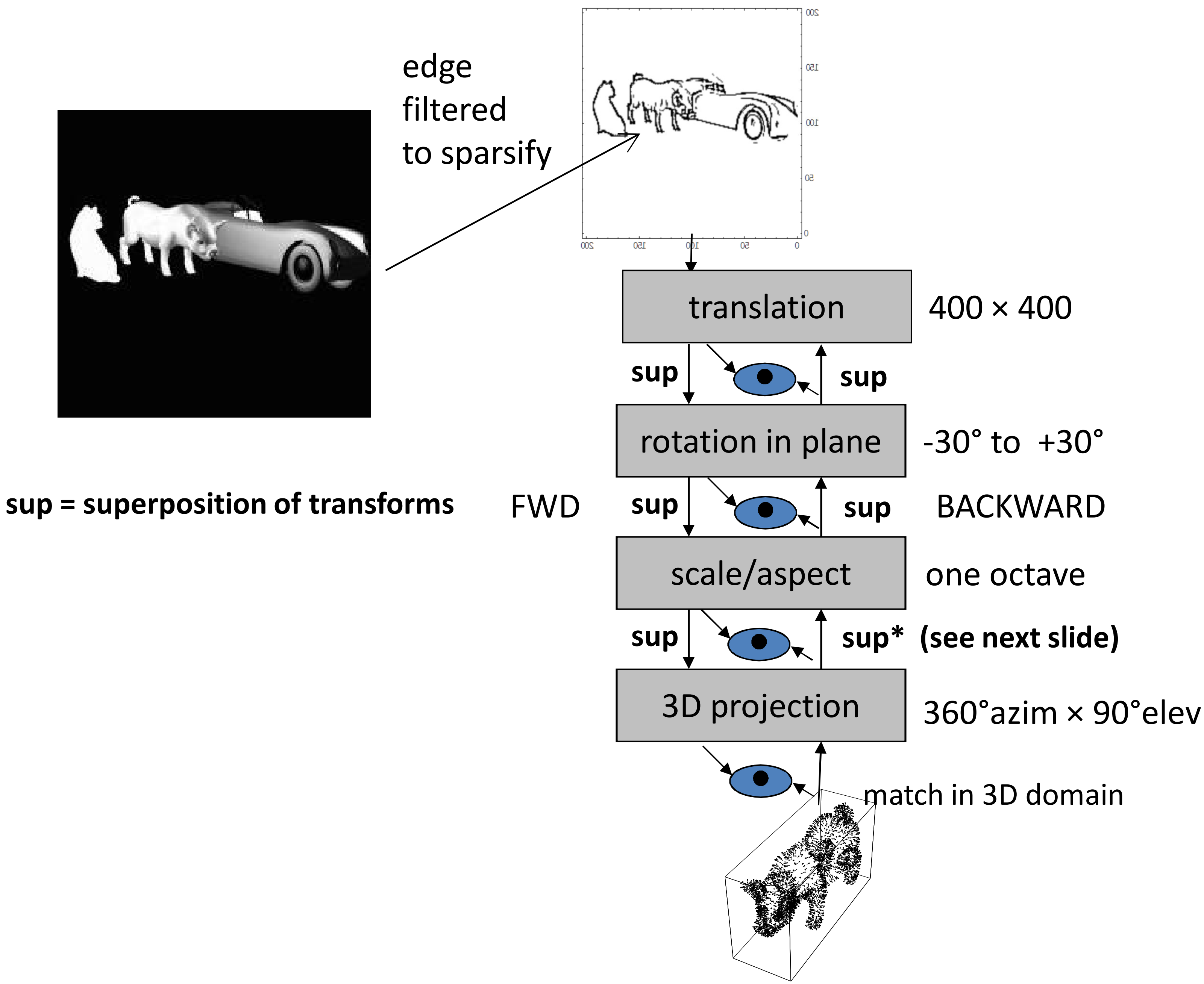
$$\Delta g_i^l = -k_l \left(1 - \frac{q_i^l}{\max\left(\mathbf{q}^l\right)}\right)$$

$$\mathbf{b}^3 = \sum_k g_k^3 \cdot \mathbf{w}_k$$

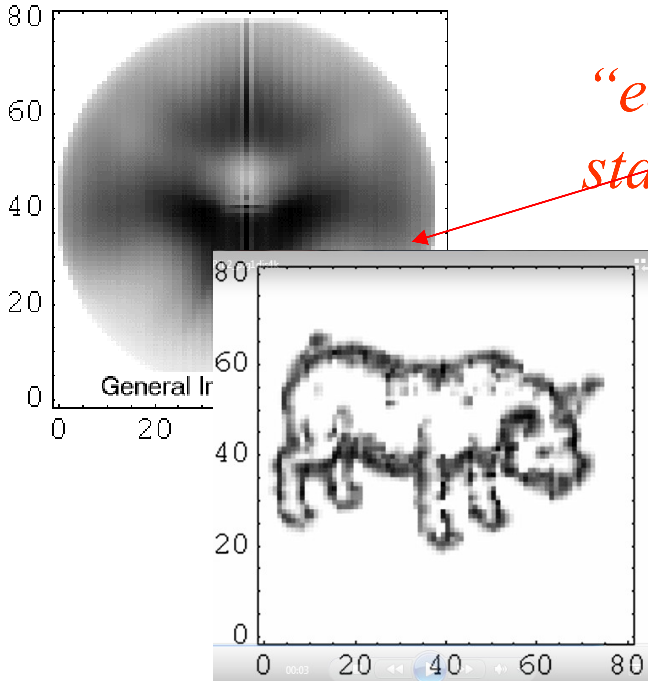
or $\mathbf{b}^3 = \mathbf{w}$

A simple MSC system: *representation* is computed OUTSIDE the MSC for both input image and models

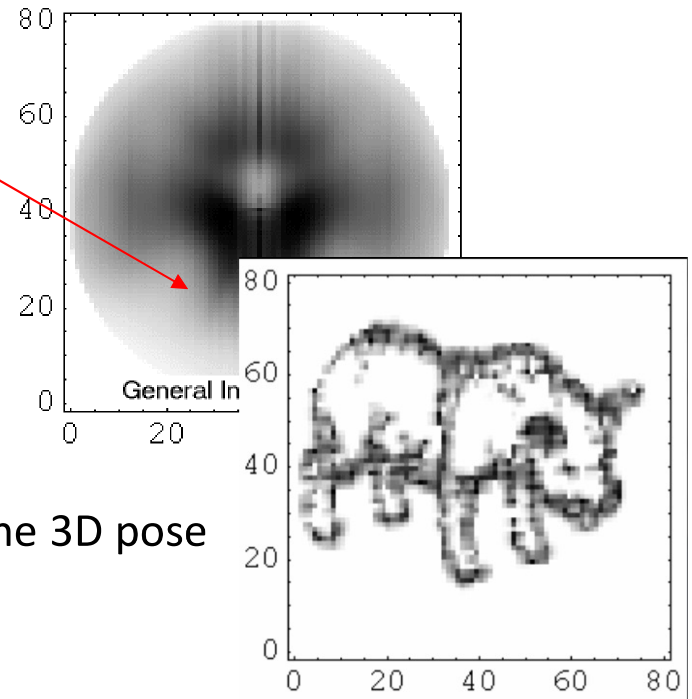



edge
filtered
to sparsify
translation
400 × 400
sup
sup
rotation in plane
-30° to +30°
sup = superposition of transforms
FWD
sup
sup
BACKWARD
scale/aspect
one octave
sup
sup* (see next slide)
3D projection
360°azim × 90°elev
match in 3D domain

two
different
views of
target

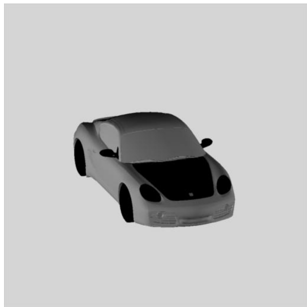


backward path
“edge plausibility states”

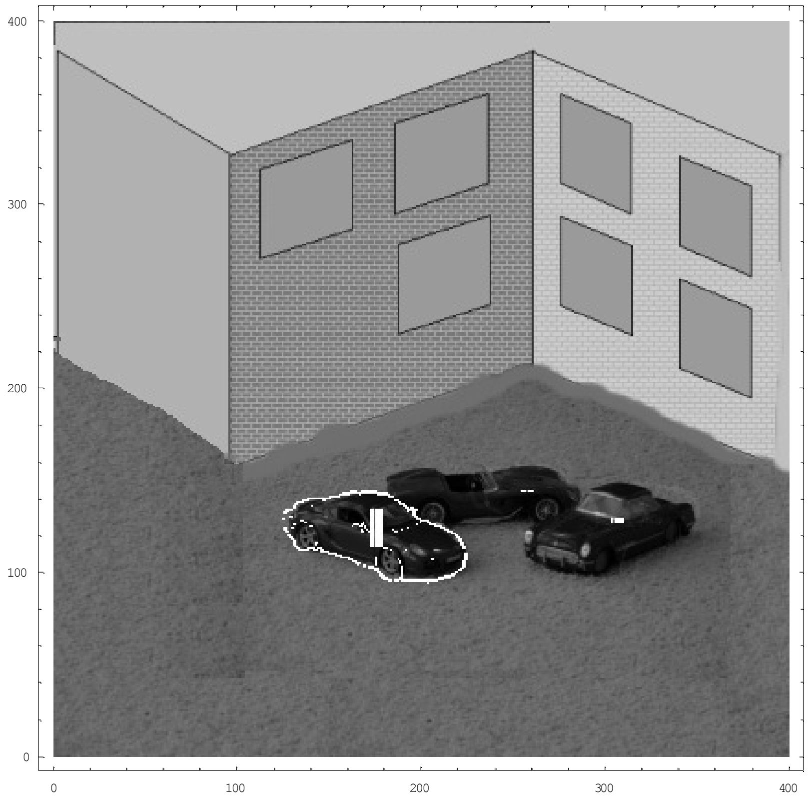


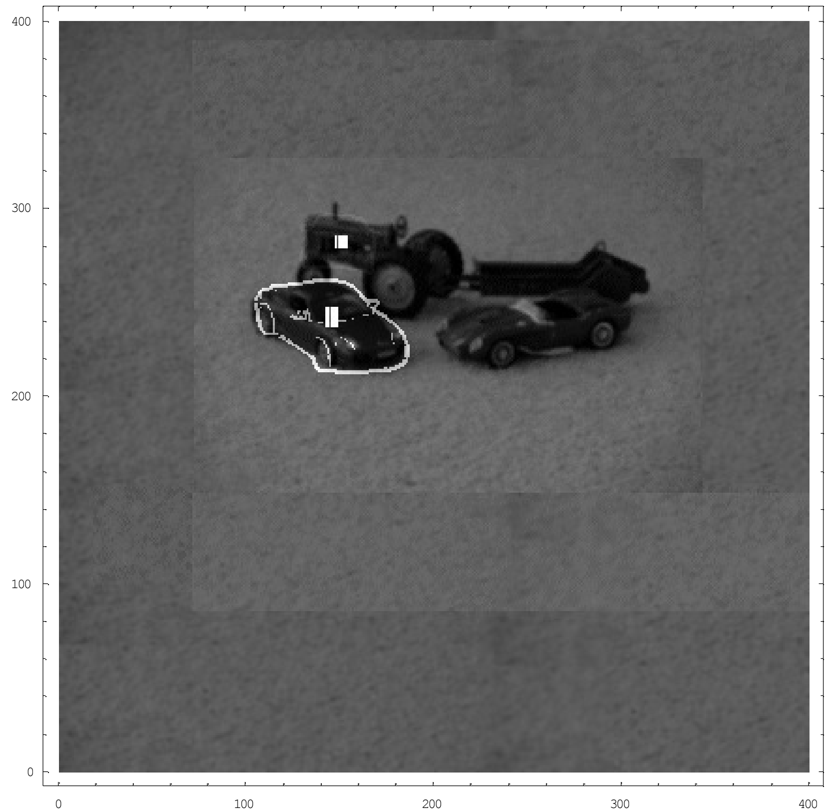
converges to determine 3D pose

Crowded targets



3D Model of Cayman





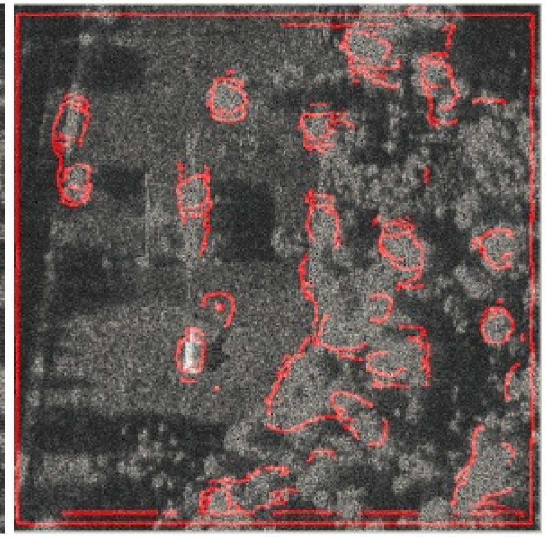
Natural Environments

Distinguishing Similar Targets

Multiple Imaging Modalities

SAR ATR (from Petersen, Murphy and Rodriguez, JHU/APL)

(a) Embedded Target



(b) Second Iteration

(c) Third Iteration

(d) Eighth Iteration

MSC is mathematically rigorous and traceable*.

Correspondence is a measure of how well an input pattern **r** matches a memory template **w** after **r** has been subjected to a sequence of transformations, t_{j1}, t_{j2}… t_{jL}.

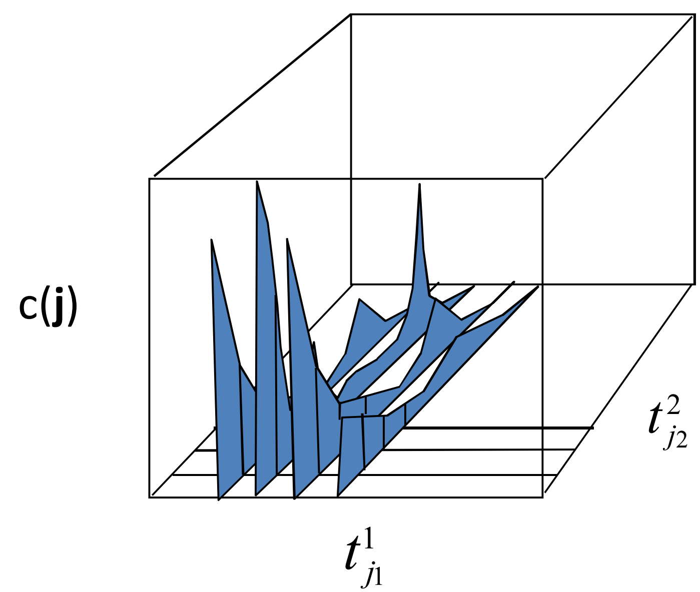
$$c(\mathbf{j}) = \left\langle t^{L}_{jL} \circ \cdots \circ t^{2}_{j2} \circ t^{1}_{j1}(\mathbf{r}), \mathbf{w} \right\rangle$$

or more compactly

$$c(\mathbf{j}) = \left\langle \mathop{\circ}_{l=1}^{L} t^{l}_{jl}(\mathbf{r}), \mathbf{w} \right\rangle$$

$$\mathbf{x} = \arg\max c(\mathbf{j}) = \left\langle \mathop{\circ}_{i=1}^{L} t^{i}_{j_i}(\mathbf{u}), \mathbf{v}^{L+1}_{j_{L+1}} \in \mathbf{V} \right\rangle$$

A ***correspondence terrain*** **C** is the array of correspondences, c(**j**), produced by all sequences of allowable transformations.



*Greg Arnold, while at AFRL/Sensors Directorate, Wright Patterson

A mathematical equivalence allows us to ***predict*** the probability of convergence to a correct solution…

$$q_j^m = \left\langle \mathop{\circ}_{l=m+1}^{L} \left(\sum_i g_i^l \cdot t_i^l \right) \circ t_j^m \circ \left(\mathop{\circ}_{l=\varnothing,1}^{m-1} \left(\sum_i g_i^l \cdot t_i^l \right) \right) (\mathbf{u}), \mathbf{v} \right\rangle$$

$$q_j^m = \sum_{\forall \mathbf{z} | z_m = j} \left(c(\mathbf{z}) \cdot \prod_{l \neq m} g_{z_l}^l \right) \quad *$$

…by computing *q*'s from correspondence terrain

* Introduced in [Harker, Gedeon, Vogel 2007, Journal of Mathematic Imaging and Vision]

Test correspondence terrains with the appropriate statistics produces a probability of correct convergence not tied to a particular set of input images.

Power	NonTargR	PopThr	Pop PerLoc	DistrLoci	Score 1	Score 2	Score 3	S%/ws%
2	0.9	0.03	2.2,3.3,3.1	18	0.94/0.94	0.84/0.86	0.84/0.84	42/46
	0.9	0.7	2.3,3.2,3.1	41	0.74/0.76	0.70/0.72	0.70/0.72	22/26
2	0.8	0.03	2.2,2.9,3.3	18	0.98/0.98	1.0/1.0	0.98/1.0	72/74
		0.07	3.6,3.0,2.6	41	0.94/0.98	0.92/0.92	0.88/0.88	30/38
		0.14	3.7,3.4,2.4	78	0.80/0.80	0.70/0.76	0.70/0.72	20/24
		0.3	3.3,2.9,3.1	158	0.44/0.48	0.68/0.70	0.50/0.54	12/16
3	0.8	0.07	2.5,3.1,3.5	41	0.94/0.94	0.94/0.94	0.94/0.94	52/58
1	0.8	0.07	2.6,3.1.3.4	41	0.80/.84	0.74/0.74	0.74/0.74	16/20
2	0.7	0.03	2.1,2.4,2.5	18	1.0/1.0	1.0/1.0	1.0/1.0	76/80
		0.07	2.9,3.9,3.2	41	1.0/1.0	1.0/1.0	0.96/0.96	46/52
		0.14	2.8,3.6,3.0	78	0.96/0.96	0.92/0.94	0.94/0.94	38/44
		0.3	3.1,2.8,2.8	158	0.68/0.72	0.84/0.86,	0.58/0.70	10/16
3	0.7	0.14	1.9,3.3,2.5	78	0.92/0.92	0.96/0.96	0.94/0.96	40/44
1	0.7	0.14	2.7,3.2,2.9	78	0.84/0.86	0.86/0.90	0.80/0.86	22/24
2	0.6	0.03	2.5,1.8,3.7	18	1.0/1.0	1.0/1.0	1.0/1.0	88/90
		0.07	3.3,3.5,3.2	41	1.0/1.0	0.98/1.0	0.98/0.98	56/62
		0.14	2.8,2.4,3.3	78	1.0/1.0	0.96/0.98	1.0/1.0	36/46
		0.3	3.4,3.4,3.2	158	0.92/0.96	0.98/0.98	0.98/0.98	24/34
		0.6	3.5,3.7,3.6	274	0.88/0.92	0.76/0.78	0.78/0.86	18/20
2	0,5	0.3	3.1,3.3,3.3	158	0.94/0.96	1.0/1.0	0.98/1.0	42/52

Yellow > ~90% correct solutions.

Handles highly cluttered environments…

(VIVID dataset V3V300009_009.avi -- frames 40-750, every tenth frame)

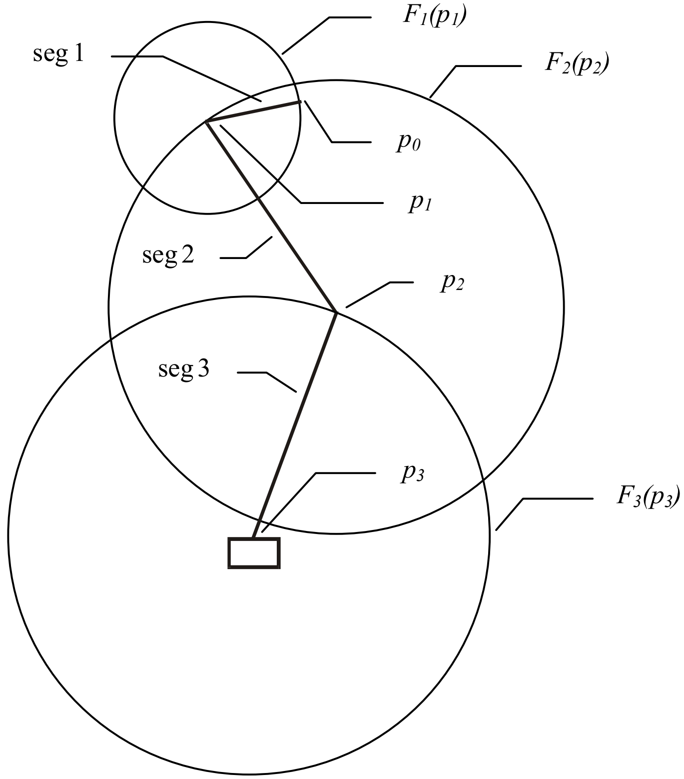
Performance on VIVID dataset, all frames independent

	Correct	Alignment Error	Hi Con False Pos	Lo Con False Pos
Non-Occluded	18/18 (100%)	0 (0%)	0 (0%)	0 (0%)
Small Occlusion	32/36 (89%)	2 (5.5%)	1 (2.7%)	1 (2.7%)
Large Occlusion	6/16 (37.5%)	6 (37.5%)	3 (18.8%)	1 (6.25%)

Summary

	Correct	Alignment Error	Total False Pos
No or Small Occlusion	50/54 (**92.6%**)	2 (**3.7%**)	2 (**3.7%**)
Large Occlusion	6/16 (37.5%)	6 (37.5%)	4 (25%)

Another inverse problem with decomposable transformations: Inverse kinematics (with obstacles)



armx	*armx*	*targx*	*targy*
20	40	64	58

segment	*θ min*	*θ max*	*θ incr*	*radius*	*θ relmin*	*θ relmax*
1:	-175.0	180.0	5.0	9	-175.0	180.0
2:	-175.0	180.0	5.0	23	-175.0	180.0
3:	-80.0	80.0	5.0	31		

Case 1: Obstacle at <x,y> =<45 to 55, 30 to 50>

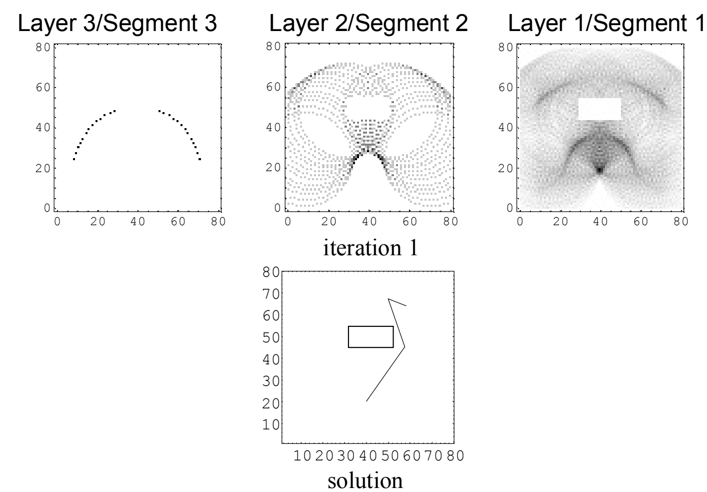
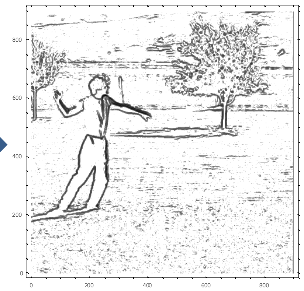


Figure 4-7. IK test including obstacles and constraints, case 1.

Using MSC to Solve for Generative Model Parameters of Articulated, Morphable Objects

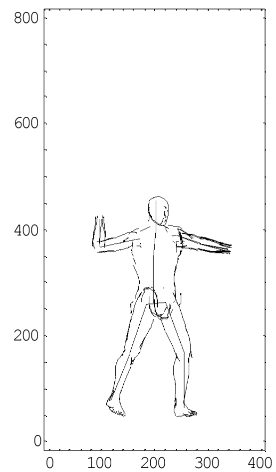


Two Concurrently Converged MSCs
Operating in Different Domains,
Linked by *Recharacterization* of Maps

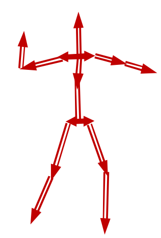
Visual
MSC

Kinematic
MSC

Morph Shell
To 2D Figure



Recharacterize
Segment Maps
to Shell



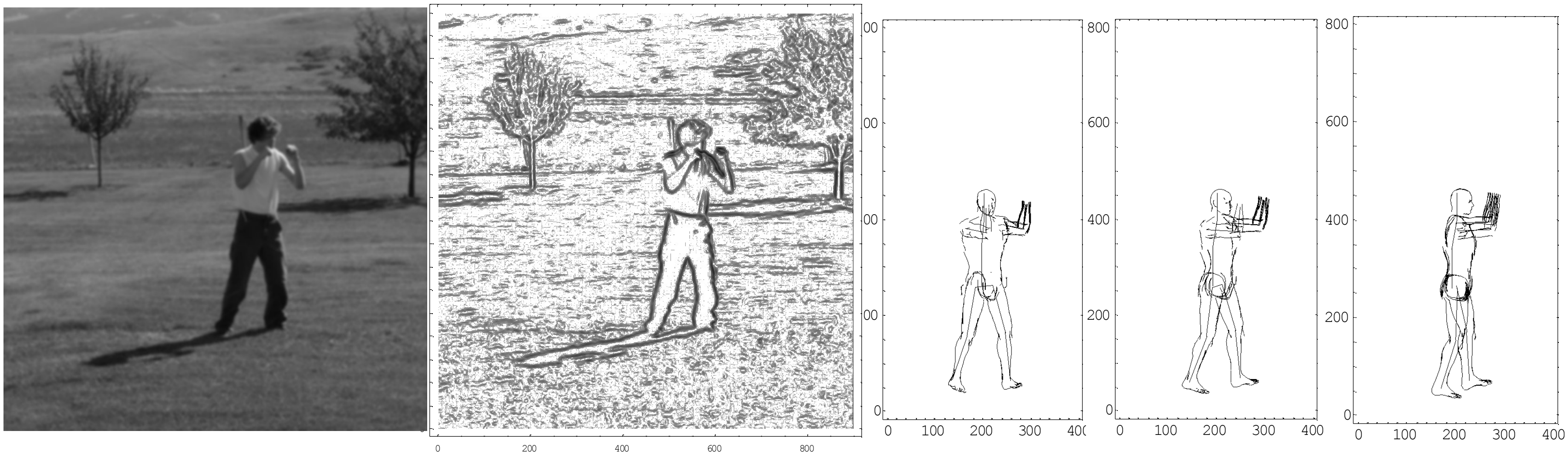
(Ref: NIPS 18)

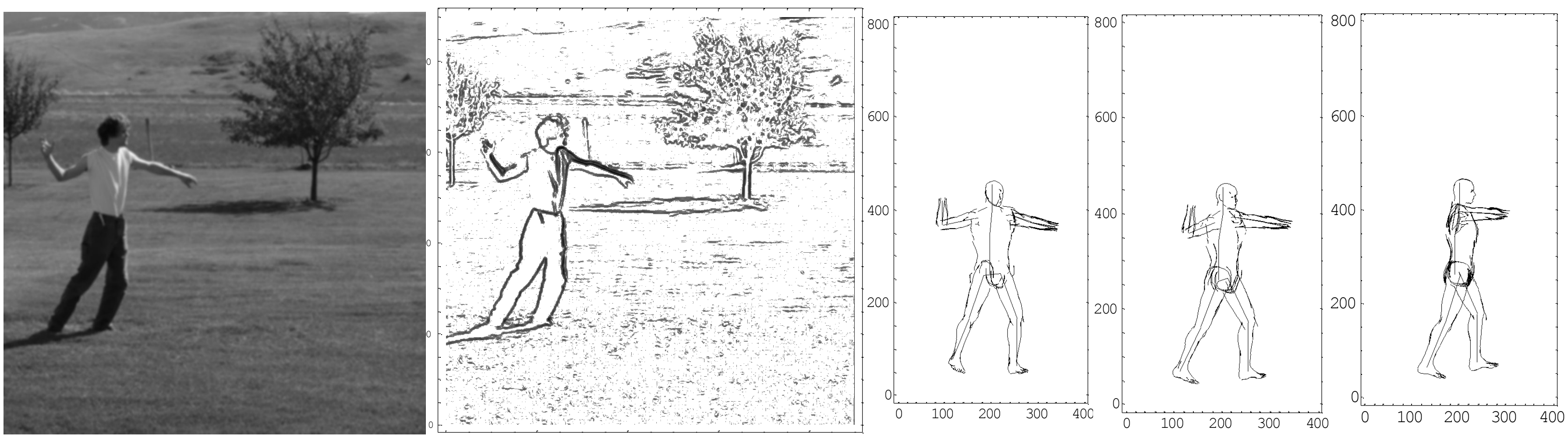
In 2D Image
Frame Ref.

In 3D Body
Frame Ref.

AFRL/Sensors SBIR Ph 1&2

Object transformations: articulations and morphings…



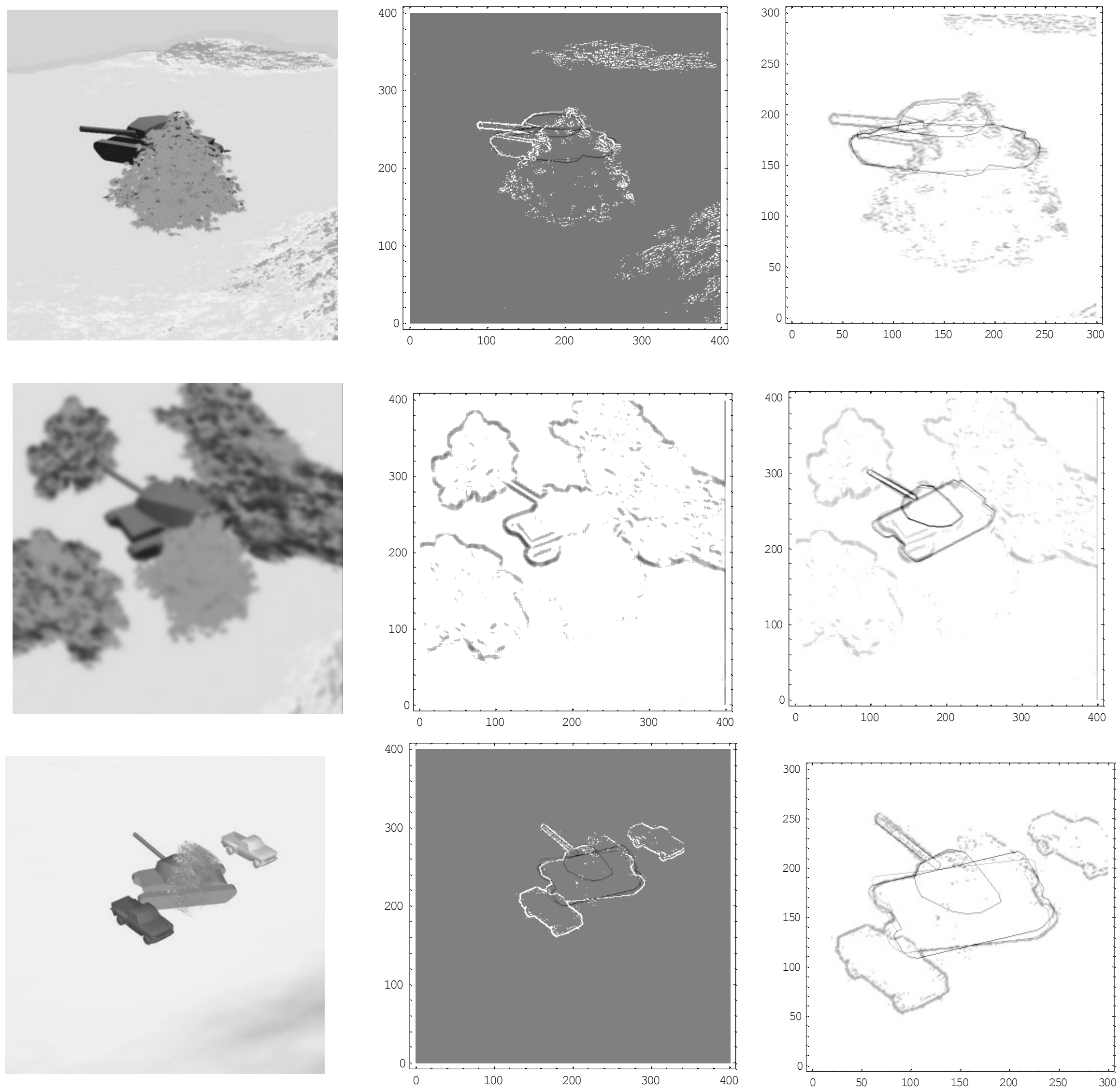


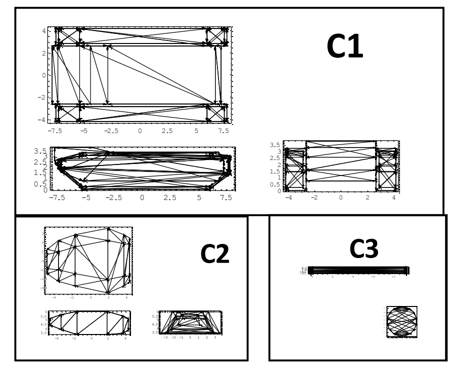
AFRL/Sensors SBIR Ph 1&2

GENERAL INTELLIGENCE

PROPRIETARY

Method Applicable to Any Articulated Object…



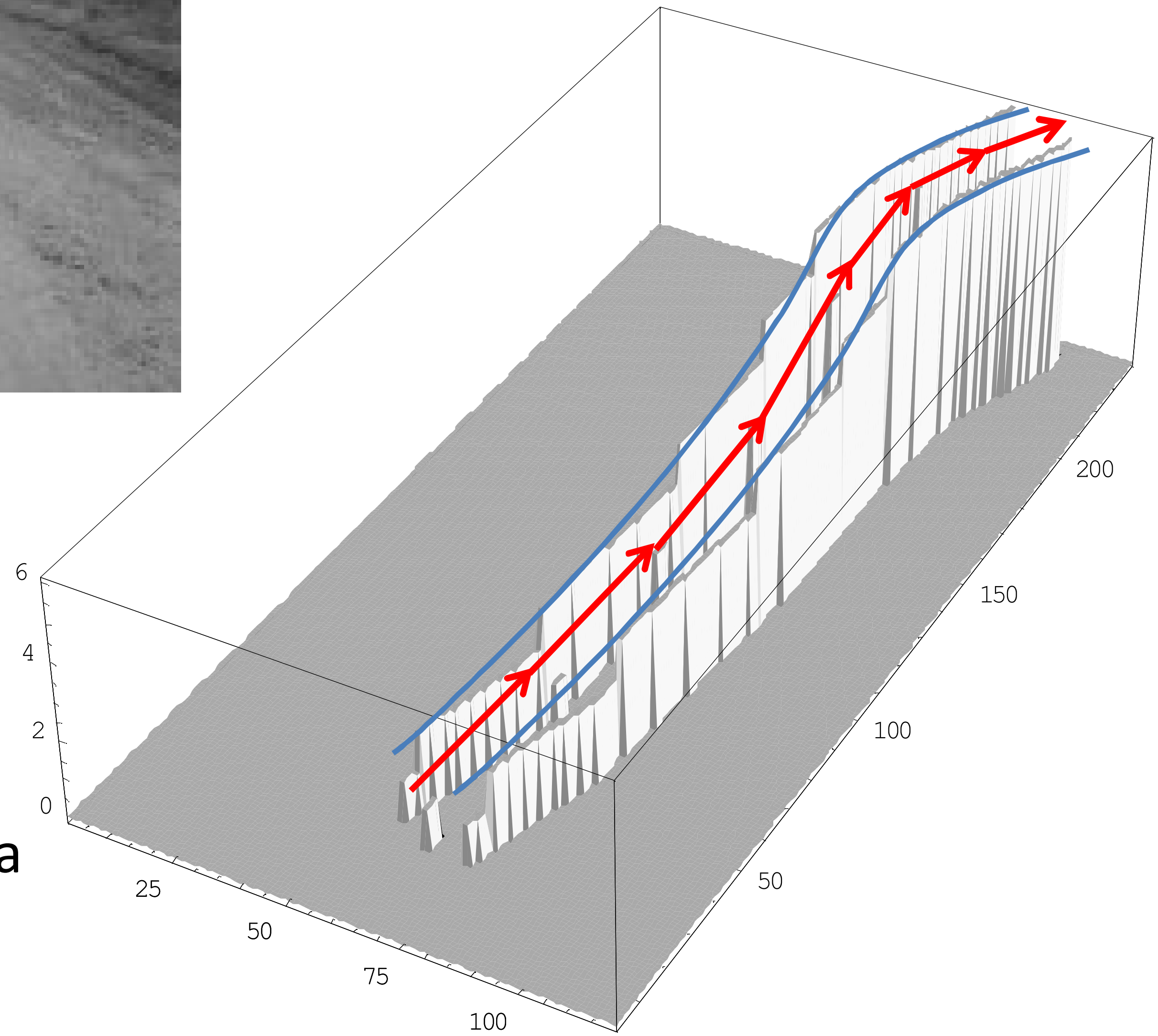


AFRL/Sensors SBIR Ph 1&2

GENERAL INTELLIGENCE

Road = an articulated plane with parallel sides.

MSC solves the road curvature in 3D from a single monocular image.



Analog Neuromorphic Implementation

Neural Code Representation Issues:

Accommodating superposition “amplitude” range

Compatibility with intra-dendritic computation

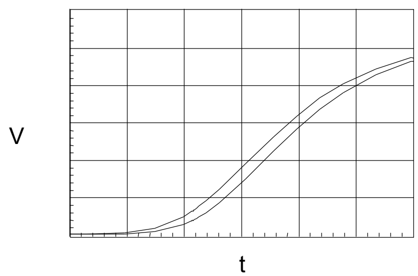
- Long pulse interaction
- Multiplicative interactions

Compatibility with basic operations

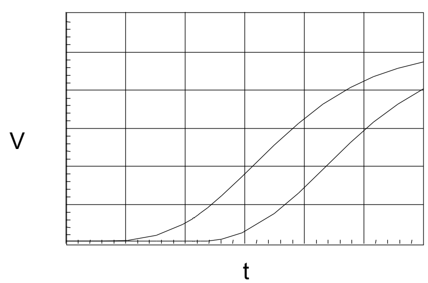
- Combination (pseudo-addition)
- Gating (pseudo-multiplication)
- Mapping
- Inhibition (pseudo-division)

Phase/Temporal Coding

... or How I Learned to Stop Worrying and Love Race Conditions

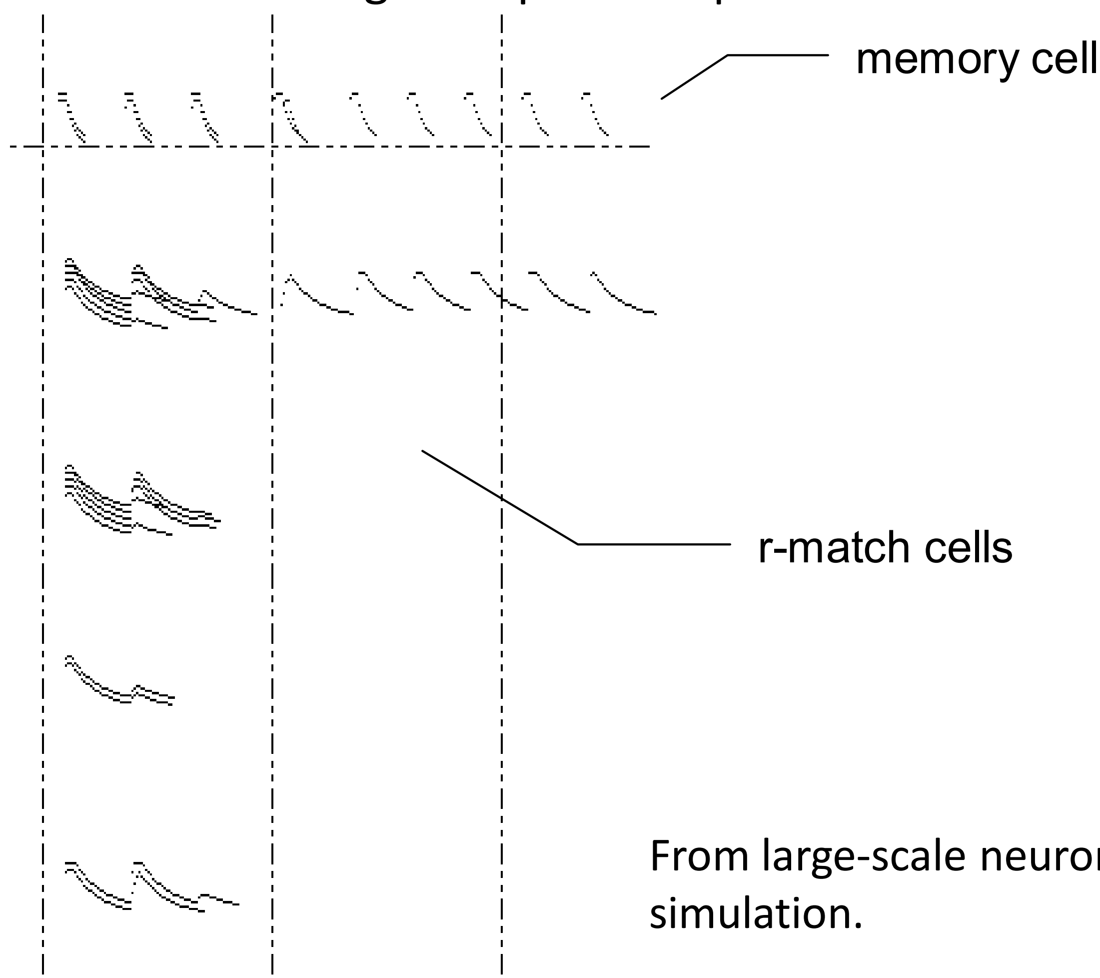


DC 1.0 vs. DC 0.7



DC 1.0 vs. DC 0.5

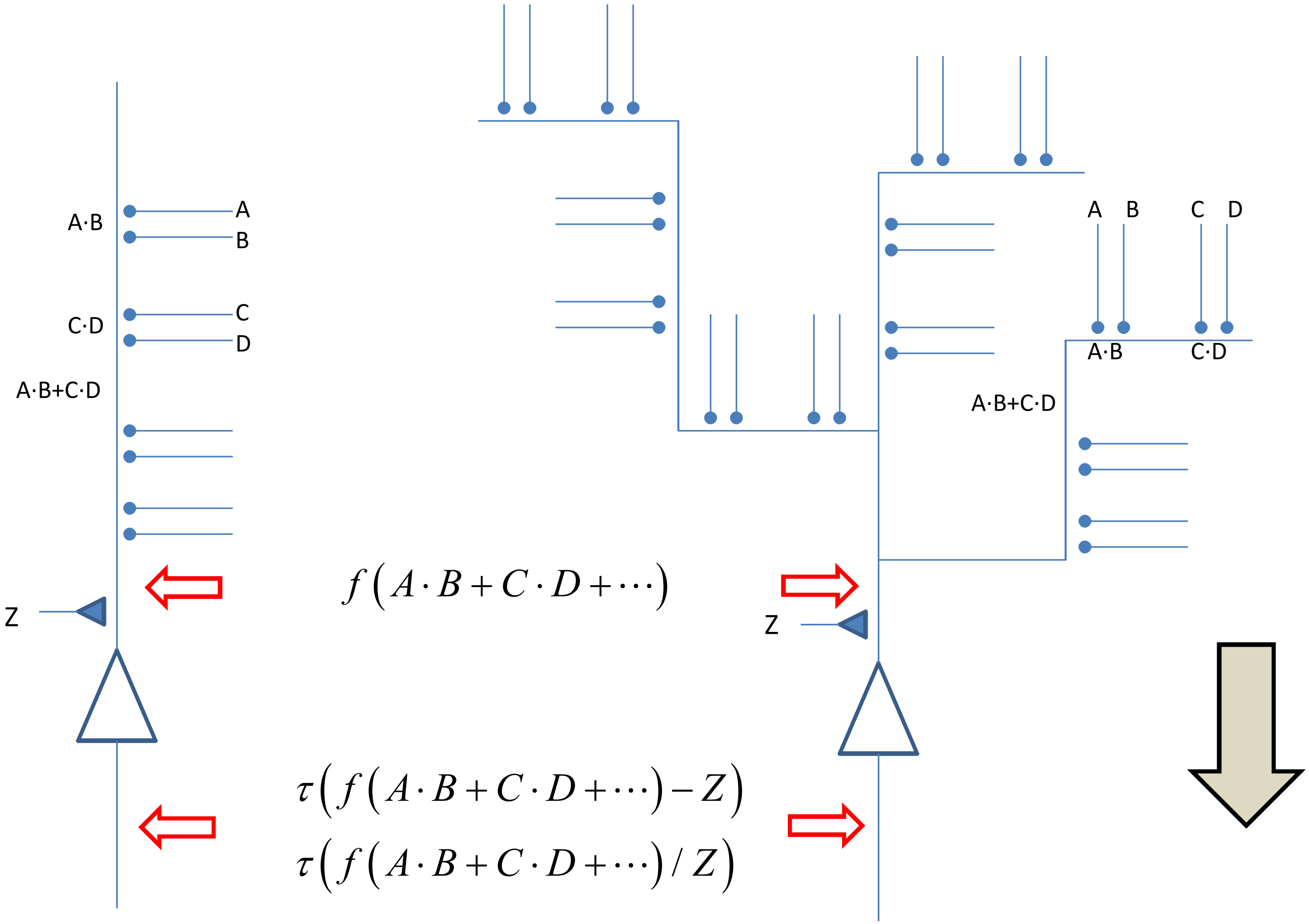
Self-normalizing: first-past-the-post



From large-scale neuronal simulation.

(*From "Map-Seeking Circuits in Visual Cognition", Chapter 5*)

Neuronal/Neuromorphic Cell Element



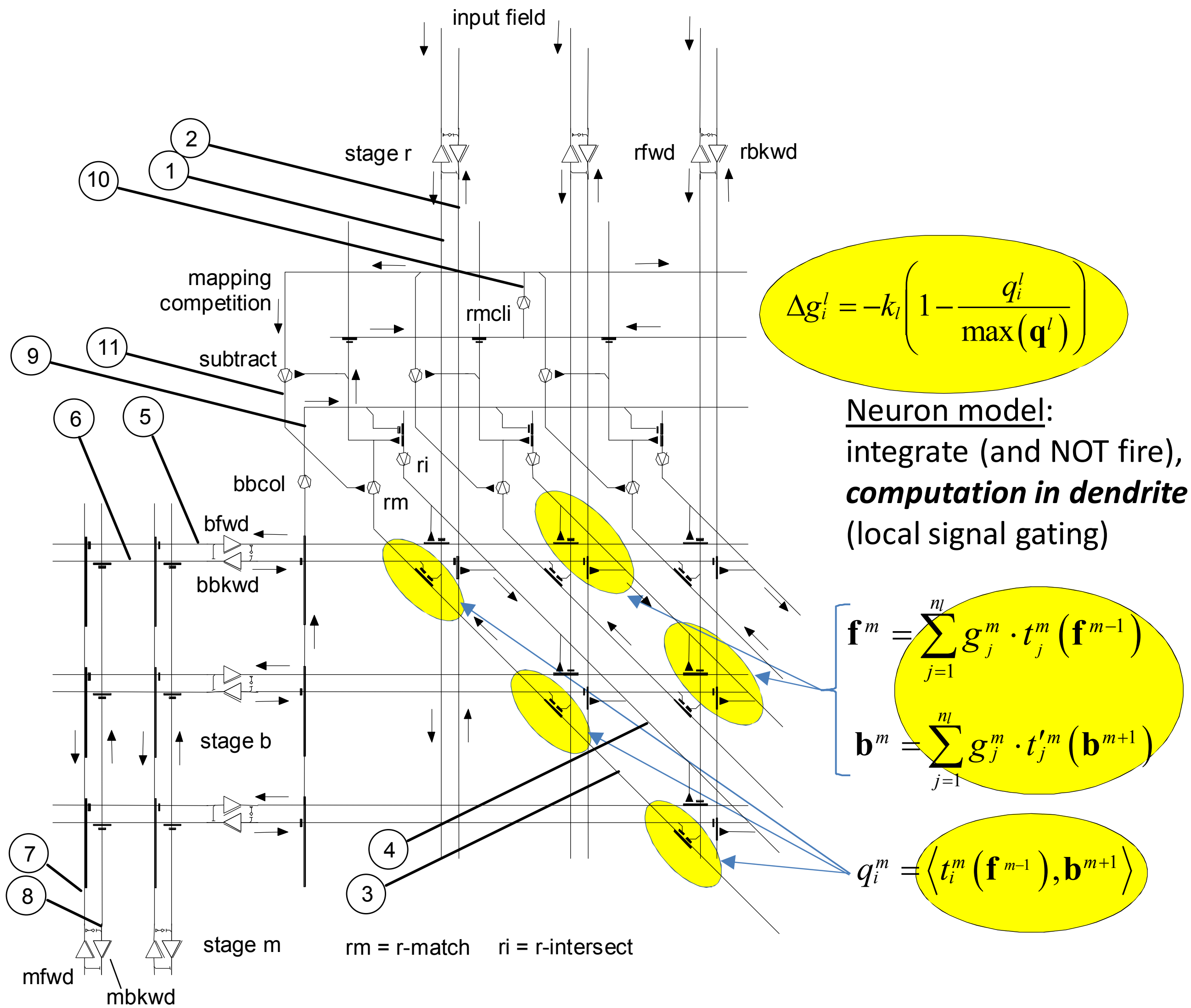
Neuronal/Neuromorphic Circuit (= Chip Layout)

Equivalence to algorithm

rm cell = *q*
ri cell = 1/*g*

rfwd = $\mathbf{f}^{L-1}$ vector
rfwd = $\mathbf{f}^{L}$ vector
bbkwd = $\mathbf{b}^{L+1}$ vector
rbkwd = $\mathbf{b}^{L}$ vector

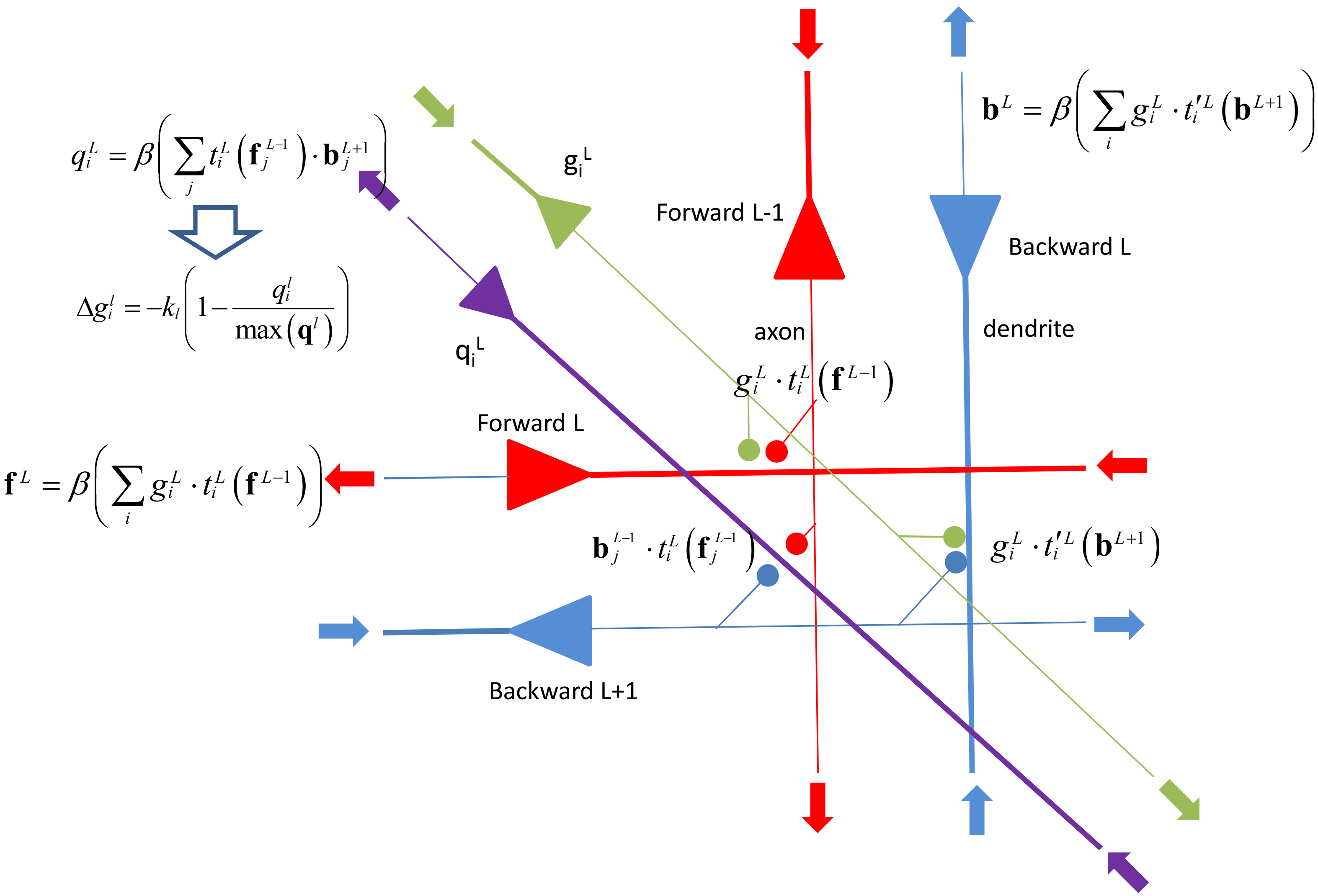
Algorithm is platonic version of this circuit



Neuron model:
integrate (and NOT fire),
computation in dendrite
(local signal gating)

(*From "Map-Seeking Circuits in Visual Cognition", Chapter 5… note nomenclature change*)

Circuit Detail



None of these synapses have memory.

Neuronal "Expense"

1 neuron per F and B vector element
4 neurons per mapping per layer
4 neurons per memory template

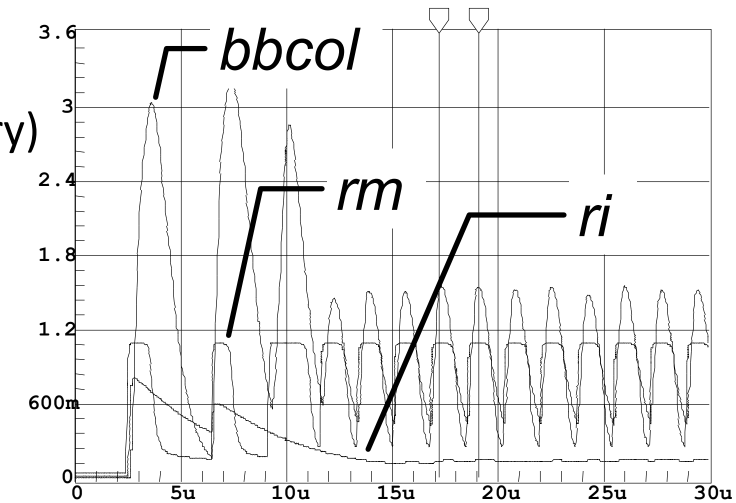
Assuming dimensions of demo MSC (ATR examples)

Approx 750K neurons for single channel (edge only)
Approx 1.9M neurons for 4 channel (e.g. 4 orientations)

Analog CMOS model Dynamics

(T-Spice, AMI 1µ device library)

Convergence to recognition state

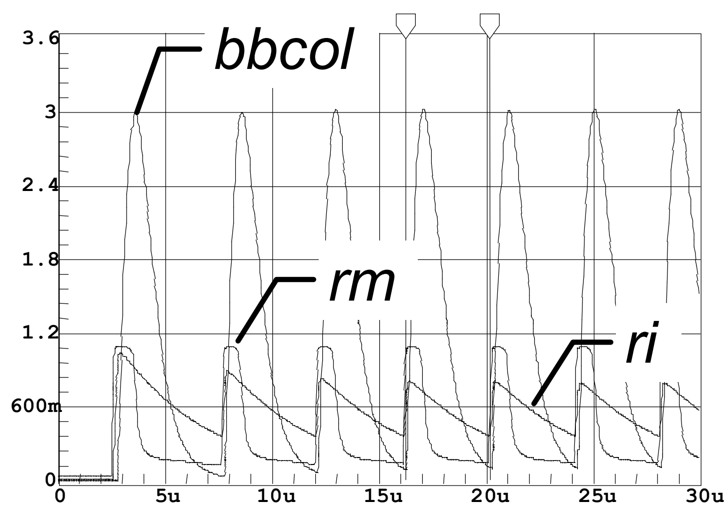


Equivalence to algorithm

rm cell = q
ri cell = 1/g

Approx 0.5MHz

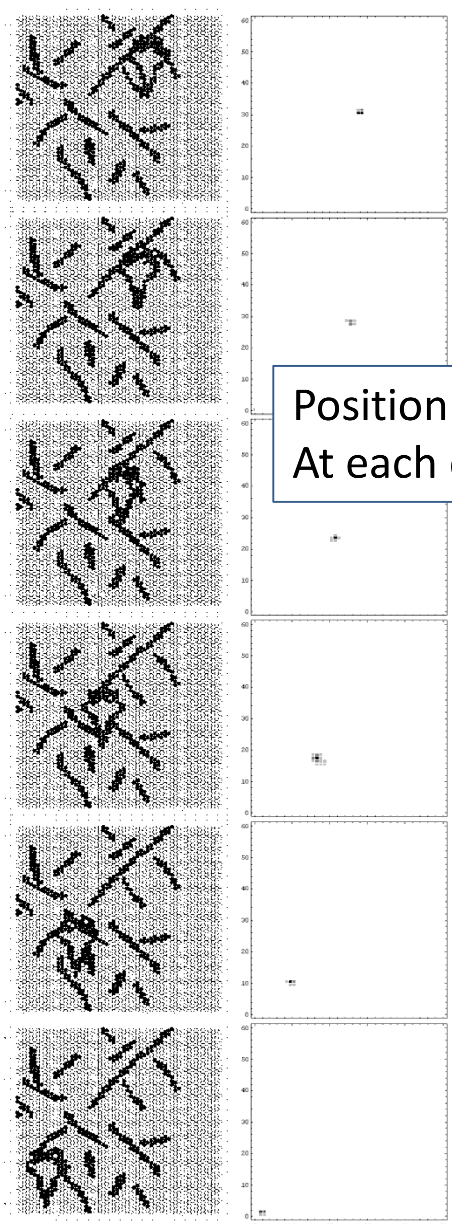
Convergence to non-recognition state



Propagation not biologically realistic.

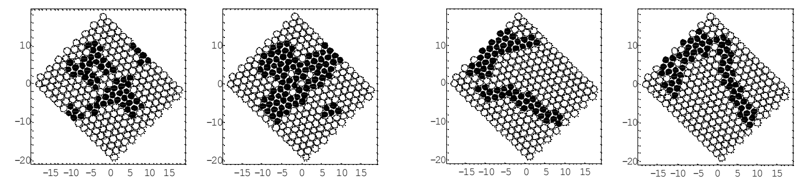
(*From "Map-Seeking Circuits in Visual Cognition", Chapter 5… note nomenclature change*)

Neuronal Circuit Tracking



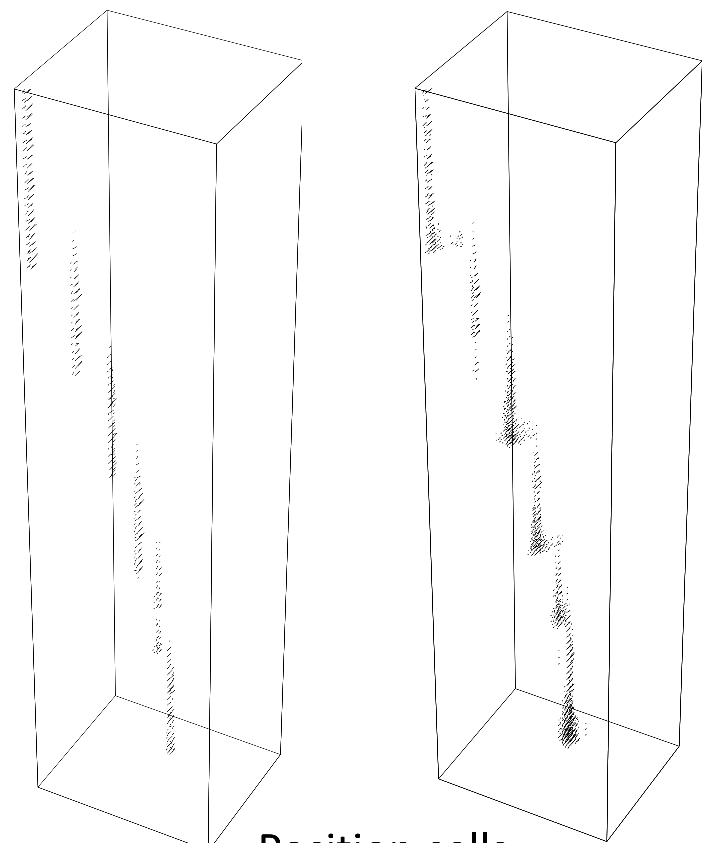
t

Position cells
At each convergence



Target model 1 Target model 2

Every
15 degree
rotation



t

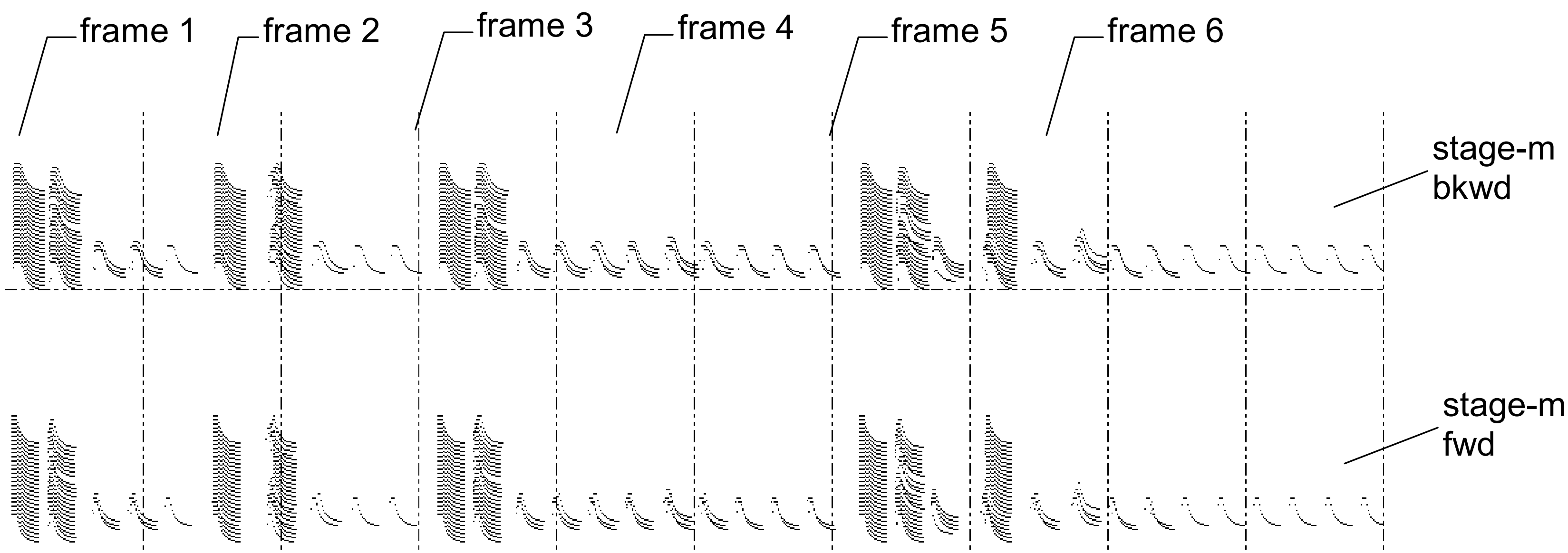
Position cells

Without clutter With clutter

(*From "Map-Seeking Circuits in Visual Cognition", Chapter 6*)

Memory cell activity during tracking (25 memory cells)

Frame change asynchronous to oscillatory period



(*From "Map-Seeking Circuits in Visual Cognition", Chapter 6*)

MSC Summary:

-- Working, efficient, tested algorithm for object recognition.
-- Solves transformation discovery problem (“correspondence problem”).
-- Solves low resolution recognition problem.
-- Verified by multiple independent implementations (APL, Lockheed, UC Berkeley, other). In production service.

-- Inherently massively fine-grain parallel.

-- Existing basic elements for analog implementation.

-- Known engineering path to neuromorphic realization.

-- No new “science” to be discovered.

Introduction,
technical,
implementation,
references,
technical reports
available at

www.giclab.com

CIRCUIT

Arathorn, DW, 2002, *Map-Seeking Circuits in Visual Cognition, A Computational Mechanism for Biological and Machine Vision*, Stanford University Press.

MATH

Harker S, Vogel CR, Gedeon T, 2007, Analysis of Constrained Optimization Variants of the Map-Seeking Circuit Algorithm, *Journal of Mathematical Imaging and Vision*, 29, August

Gedeon T, Arathorn DW, 2007, Convergence of Map-Seeking Circuits, *Journal of Mathematical Imaging and Vision.*