Segmentation of Objects in Images and Videos

Alexandre X. Falcão and Thiago V. Spina

Institute of Computing - University of Campinas

tvspina@ic.unicamp.br

Introduction

What is segmentation of images and videos?

- The segmentation of objects in images and videos aims at extracting an object of interest from the background.

Introduction

What is segmentation of images and videos?

- The segmentation of objects in images and videos aims at extracting an object of interest from the background.
- But what is an object?

Introduction

What is segmentation of images and videos?

- The segmentation of objects in images and videos aims at extracting an object of interest from the background.
- But what is an object?

- And what is the background? $\rightarrow$ Highly dependent on the context.

Introduction

For completeness, segmentation methods should

Introduction

For completeness, segmentation methods should

- acquire and model object information;

Introduction

For completeness, segmentation methods should

- acquire and model object information;
- enhance regions wherein image properties are similar to those of the object, background, or their transition;

Introduction

For completeness, segmentation methods should

- acquire and model object information;
- enhance regions wherein image properties are similar to those of the object, background, or their transition;
- locate the object and delineate its spatial extent in the image.

Introduction

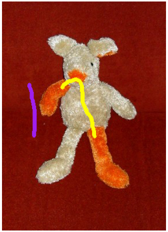
- Humans are more accurate than computers for object location.

Introduction

- Humans are more accurate than computers for object location.
- Computers are more precise than humans in object delineation.

Introduction

- Humans are more accurate than computers for object location.
- Computers are more precise than humans in object delineation.



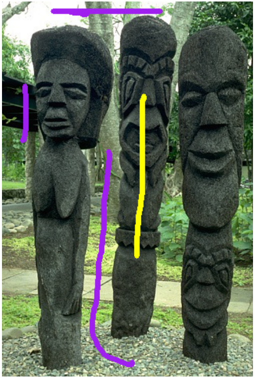
Introduction

- Humans are more accurate than computers for object location.
- Computers are more precise than humans in object delineation.

In this example, delineation was solved by object enhancement followed by binarization (i.e., a spel classification), with no need for connectivity.

When do we need connectivity?

Simple connectivity is needed for delineation when object and other disconnected parts of the background have similar properties.



When do we need connectivity?

Simple connectivity is needed for delineation when object and other disconnected parts of the background have similar properties.

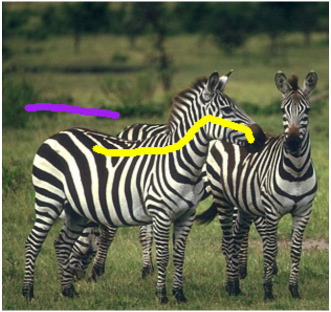
When do we need connectivity?

Simple connectivity is needed for delineation when object and other disconnected parts of the background have similar properties.

However, when do we need optimum connectivity?

When do we need optimum connectivity?

Optimum connectivity is needed for delineation when object and parts of the background with similar properties are connected to each other.

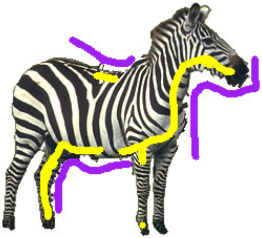


When do we need optimum connectivity?

Optimum connectivity is needed for delineation when object and parts of the background with similar properties are connected to each other.

When do we need optimum connectivity?

Optimum connectivity is needed for delineation when object and parts of the background with similar properties are connected to each other.



In this case, however, some markers needed to disconnect the object are not suitable for object enhancement.

How can we effectively address segmentation?

- Markers/object models can be used for object location and enhancement.

How can we effectively address segmentation?

- Markers/object models can be used for object location and enhancement.
- Enhancement must be intelligent to extract suitable object information.

How can we effectively address segmentation?

- Markers/object models can be used for object location and enhancement.
- Enhancement must be intelligent to extract suitable object information.
- For interactive segmentation, we can exploit a synergism between object location/correction by a human operator and computer delineation.

How can we effectively address segmentation?

- Markers/object models can be used for object location and enhancement.
- Enhancement must be intelligent to extract suitable object information.
- For interactive segmentation, we can exploit a synergism between object location/correction by a human operator and computer delineation.
- For automatic segmentation, we can exploit a synergism between object location by some object model and computer delineation.

How can we effectively address segmentation?

- Markers/object models can be used for object location and enhancement.
- Enhancement must be intelligent to extract suitable object information.
- For interactive segmentation, we can exploit a synergism between object location/correction by a human operator and computer delineation.
- For automatic segmentation, we can exploit a synergism between object location by some object model and computer delineation.
- In both cases, delineation based on optimum connectivity can be used in the image domain and/or in the feature space by simple choice of the adjacency relation.

Organization of this lecture

We will focus on

Organization of this lecture

We will focus on

- Object delineation using the image foresting transform: boundary-based, region-based, and hybrid approaches.

Organization of this lecture

We will focus on

- Object delineation using the image foresting transform: boundary-based, region-based, and hybrid approaches.
- A comparative analysis between the IFT and the min-cut/max-flow algorithms for region-based segmentation.

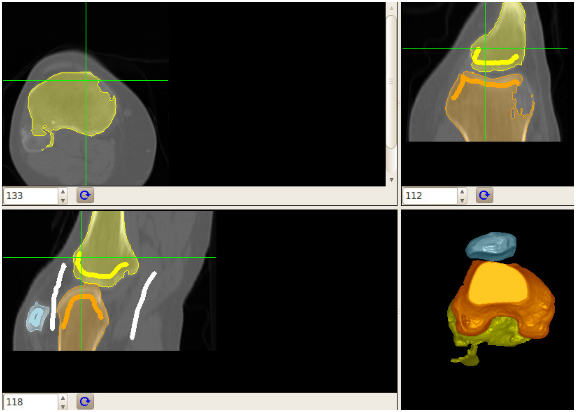
Organization of this lecture

We will focus on

- Object delineation using the image foresting transform: boundary-based, region-based, and hybrid approaches.
- A comparative analysis between the IFT and the min-cut/max-flow algorithms for region-based segmentation.
- Fuzzy object models and video segmentation.

Region-based object delineation

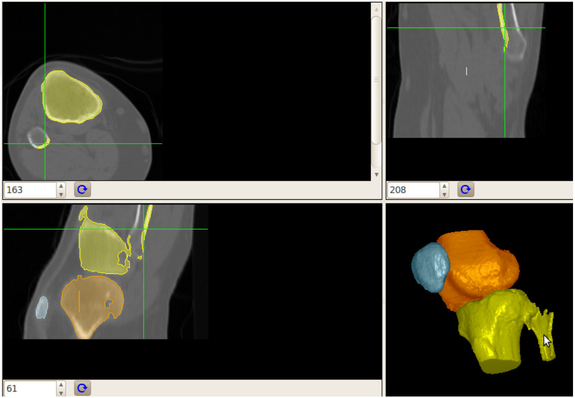
- Multiple objects can be segmented with interactive response time to the user's actions by using the differential IFT with seed competition (IFTSC).



- Interactive 3D visualization is crucial to help on object location and correction.

Region-based object delineation

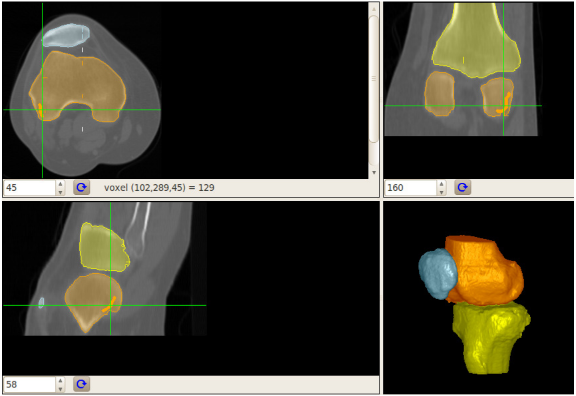
- Multiple objects can be segmented with interactive response time to the user's actions by using the differential IFT with seed competition (IFTSC).



- Interactive 3D visualization is crucial to help on object location and correction.

Region-based object delineation

- Multiple objects can be segmented with interactive response time to the user's actions by using the differential IFT with seed competition (IFTSC).



- Interactive 3D visualization is crucial to help on object location and correction.

Differential IFT with seed competition (IFTSC)

- In this case, the object is an optimum-path forest for $f_{\max}$ rooted at internal seeds.

$$
\begin{aligned}
f_{\max}(\langle t \rangle) &= \begin{cases} 0 & \text{if } t \in \mathcal{S} = \mathcal{S}_i \cup \mathcal{S}_e \\ +\infty & \text{otherwise} \end{cases} \\
f_{\max}(\pi_s \cdot \langle s, t \rangle) &= \max\{f_{\max}(\pi_s), w(s,t)\},
\end{aligned}
$$

where $\mathcal{S}_i$ and $\mathcal{S}_e$ are internal and external seed sets.

Differential IFT with seed competition (IFTSC)

- In this case, the object is an optimum-path forest for $f_{\max}$ rooted at internal seeds.

$$
\begin{aligned}
f_{\max}(\langle t\rangle) &= \begin{cases} 0 & \text{if } t \in \mathcal{S} = \mathcal{S}_i \cup \mathcal{S}_e \\ +\infty & \text{otherwise} \end{cases} \\
f_{\max}(\pi_s \cdot \langle s, t\rangle) &= \max\{f_{\max}(\pi_s), w(s,t)\},
\end{aligned}
$$

where $\mathcal{S}_i$ and $\mathcal{S}_e$ are internal and external seed sets.

- The dual formulation holds for fuzzy connected segmentation.

Differential IFT with f_{max} and seed competition (IFTSC)

Algorithm

– DIFTSC - ALGORITHM

1. $(V, P, R, \mathcal{F}) \leftarrow$ DIFT-FORESTREMOVAL$(V, P, R, \mathcal{A}, \mathcal{R}_{\mathcal{M}})$.
2. $\mathcal{F} \leftarrow \mathcal{F} \setminus \mathcal{S}$.
3. **While** $\mathcal{S} \neq \emptyset$, remove t from $\mathcal{S}$, set $V(t) \leftarrow 0$,
4. set $L(t) \leftarrow \lambda(t)$, $R(t) \leftarrow t$, $P(t) \leftarrow nil$, and $\mathcal{F} \leftarrow \mathcal{F} \cup \{t\}$.
5. **While** $\mathcal{F} \neq \emptyset$, remove t from $\mathcal{F}$ and insert t in Q.
6. **While** Q is not empty **do**
7. Remove s from Q such that $V(s)$ is minimum.
8. **For each** $t \in \mathcal{A}(s)$, **do**
9. Compute $tmp \leftarrow \max\{V(s), w(s,t)\}$.
10. **If** $tmp < V(t)$ or $P(t) = s$, **then**
11. **If** $t \in Q$, **then** remove t from Q.
12. Set $P(t) \leftarrow s$, $V(t) \leftarrow tmp$, $R(t) \leftarrow R(s)$,
13. $L(t) \leftarrow L(s)$, and Insert t in Q.

Differential IFT with f_{max} and seed competition (IFTSC)

Algorithm

– DIFT-FORESTREMOVAL

INPUT: *Maps V, P, R, adjacency $\mathcal{A}$, and set $\mathcal{R}_{\mathcal{M}}$ of selected roots.*
OUTPUT: *Maps V, P, and set $\mathcal{F}$ of frontier pixels.*
AUXILIARY: *FIFO Queue T.*

1. *Set $\mathcal{F} \leftarrow \emptyset$.*
2. **For each** $t \in \mathcal{R}_{\mathcal{M}}$, **do** *insert t in T, set $V(t) \leftarrow +\infty$ and $P(t) \leftarrow nil$.*
3. **While** $T \neq \emptyset$, **do**
4. *Remove s from T.*
5. **For each** $t \in \mathcal{A}(s)$, **do**
6. **If** $P(t) = s$, **then**
7. *Set $V(t) \leftarrow +\infty$, $P(t) \leftarrow nil$ and insert t in T.*
8. **Else If** $R(t) \notin \mathcal{R}_{\mathcal{M}}$, **then** $\mathcal{F} \leftarrow \mathcal{F} \cup \{t\}$.
9. *Set $\mathcal{R}_{\mathcal{M}} \leftarrow \emptyset$.*

Boundary-based object delineation

An ordered sequence of optimum paths can define the object's boundary by several different ways.

Boundary-based object delineation

An ordered sequence of optimum paths can define the object's boundary by several different ways.

- Methods, such as live-wire and riverbed, present the paths as the user selects boundary points and moves the cursor on the image.

Boundary-based object delineation

An ordered sequence of optimum paths can define the object's boundary by several different ways.

- Methods, such as live-wire and riverbed, present the paths as the user selects boundary points and moves the cursor on the image.
- Iterative live-wire uses, as input, points nearby the boundary and executes live-wire several times, replacing those points by the mid-segment ones until convergence.

Boundary-based object delineation

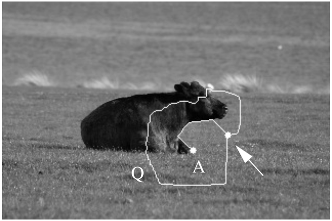
An ordered sequence of optimum paths can define the object's boundary by several different ways.

- Methods, such as live-wire and riverbed, present the paths as the user selects boundary points and moves the cursor on the image.
- Iterative live-wire uses, as input, points nearby the boundary and executes live-wire several times, replacing those points by the mid-segment ones until convergence.

They can be implemented by a sequence of IFTs using 4- or 8-adjacency relation and suitable connectivity function.

Live-wire-on-the-fly

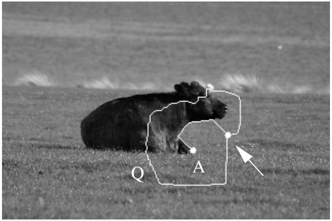
In live-wire-on-the-fly (LWOF), optimum paths are incrementally computed from the moving wavefront Q.



- The user selects a point A on the object's boundary, and

Live-wire-on-the-fly

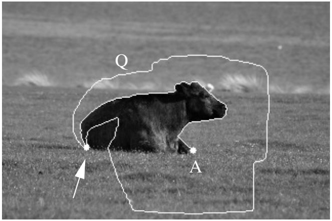
In live-wire-on-the-fly (LWOF), optimum paths are incrementally computed from the moving wavefront Q.



- The user selects a point A on the object's boundary, and
- for any subsequent position of the cursor, an optimum path from A to that position is displayed in real time.

Live-wire-on-the-fly

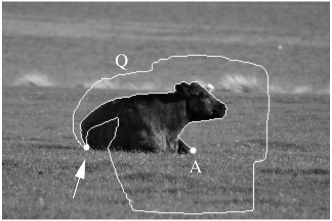
In live-wire-on-the-fly (LWOF), optimum paths are incrementally computed from the moving wavefront Q.



- The user selects a point A on the object's boundary, and
- for any subsequent position of the cursor, an optimum path from A to that position is displayed in real time.
- When the cursor is close to the boundary, the path snaps on to it.

Live-wire-on-the-fly

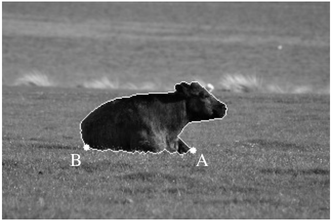
In live-wire-on-the-fly (LWOF), optimum paths are incrementally computed from the moving wavefront Q.



- The user selects a point A on the object's boundary, and
- for any subsequent position of the cursor, an optimum path from A to that position is displayed in real time.
- When the cursor is close to the boundary, the path snaps on to it.
- The user may accept it as a boundary segment, and

Live-wire-on-the-fly

In live-wire-on-the-fly (LWOF), optimum paths are incrementally computed from the moving wavefront Q.



- The user selects a point A on the object's boundary, and
- for any subsequent position of the cursor, an optimum path from A to that position is displayed in real time.
- When the cursor is close to the boundary, the path snaps on to it.
- The user may accept it as a boundary segment, and
- the process is repeated from its terminus B until the user decides to close the contour.

Live-wire-on-the-fly (LWOF)

The IFT algorithm with early termination and function f_{sum} finds optimum paths from a starting point s^* on counter-clockwise oriented boundaries.

$$
\begin{aligned}
f_{sum}^{\circlearrowleft}(\langle t \rangle) &= \begin{cases} 0 & \text{if } t = s^* \\ +\infty & \text{otherwise} \end{cases} \\
f_{sum}^{\circlearrowleft}(\pi_s \cdot \langle s, t \rangle) &= \begin{cases} f_{sum}^{\circlearrowleft}(\pi_s) + \bar{w}^{\beta}(s,t) & \text{if } O(l) \geq O(r) \\ f_{sum}^{\circlearrowleft}(\pi_s) + K^{\beta} & \text{otherwise,} \end{cases}
\end{aligned}
$$

where l and r are the spels at the left and right sides of arc $\langle s, t \rangle$. The weights $\bar{w}(s,t)$ are lower on the boundary than inside and outside it and $\beta \geq 1$ favors longer segments.

Path computation from s^* to u in LWOF

Algorithm

– PATH COMPUTATION FROM s^* TO u IN LWOF

1. **If** $V(u) = +\infty$ *or* $u \in Q$, **then**
2. *Set* $s \leftarrow nil$.
3. *While* $Q \neq \emptyset$ *and* $s \neq u$, *do*
4. *Remove from* Q *a spel* s *such that* $V(s)$ *is minimum.*
5. *For each* $t \in \mathcal{A}(s)$ *such that* $V(t) > V(s)$, *do*
6. **If** $O(l) \geq O(r)$,
7. **then** *set* $tmp \leftarrow V(s) + \bar{w}^{\beta}(s,t)$
8. **Else** *set* $tmp \leftarrow V(s) + K^{\beta}$.
9. **If** $tmp < V(t)$, **then**
10. *If* $V(t) \neq +\infty$, *remove* t *from* Q.
11. *Set* $P(t) \leftarrow s$ *and* $V(t) \leftarrow tmp$.
12. *Insert* t *in* Q.

Riverbed

Riverbed simulates the behavior of water flowing through a riverbed, always seeking lower ground levels, snaking through the river bends, instead of short-cutting the path as in live-wire. This leads to the following connectivity function for a starting seed point s^*:

$$
\begin{aligned}
f_w^{\circlearrowleft}(\langle t \rangle) &= \begin{cases} 0 & \text{if } t = s^* \\ +\infty & \text{otherwise} \end{cases} \\
f_w^{\circlearrowleft}(\pi_s \cdot \langle s, t \rangle) &= \begin{cases} \bar{w}(s,t) & \text{if } O(l) \geq O(r) \\ K & \text{otherwise.} \end{cases}
\end{aligned}
$$

where l and r are the spels at the left and right sides of arc $\langle s, t \rangle$.

Riverbed

- Although $f_w^{\circlearrowleft}$ is not smooth, it selects segments such that the maximum arc weight

$$\max_{\forall (l,r) \in \mathcal{A}', L(l)=1, L(r)=0} \bar{w}(l,r)$$

is minimum, considering all possible cuts in the dual graph $(\mathcal{N}, \mathcal{A}')$.

Riverbed

- Although $f_w^{\circlearrowleft}$ is not smooth, it selects segments such that the maximum arc weight

$$\max_{\forall (l,r) \in \mathcal{A}', L(l)=1, L(r)=0} \bar{w}(l,r)$$

  is minimum, considering all possible cuts in the dual graph $(\mathcal{N}, \mathcal{A}')$.
- This implies that IFTSC and riverbed decide for the same optimum graph cut.

Riverbed

- Although $f_w^{\circlearrowleft}$ is not smooth, it selects segments such that the maximum arc weight

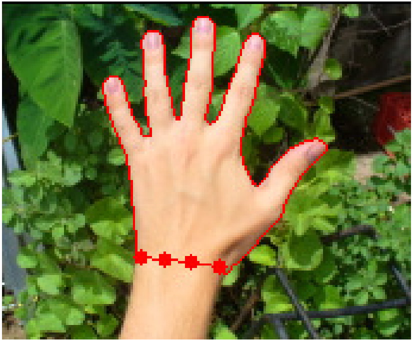
$$\max_{\forall (l,r) \in \mathcal{A}', L(l)=1, L(r)=0} \bar{w}(l,r)$$

  is minimum, considering all possible cuts in the dual graph $(\mathcal{N}, \mathcal{A}')$.
- This implies that IFTSC and riverbed decide for the same optimum graph cut.
- Riverbed is more suitable than live-wire for more intricate shapes, but live-wire can jump weakly defined parts of the boundary.

Riverbed versus live-wire

- Live-wire on complex shapes requires more anchor points.

Riverbed versus live-wire



- Live-wire on complex shapes requires more anchor points.
- Riverbed asks for more user intervention on poorly defined parts.

Riverbed versus live-wire

- Live-wire on complex shapes requires more anchor points.
- Riverbed asks for more user intervention on poorly defined parts.
- Their combination requires only two segments (live wire in cyan, riverbed in red).

Live Markers

Live Markers is another hybrid approach that takes advantage from the superior ability of LWOF on weakly defined segments and from IFTSC to handle complex 2D/3D shapes of multiple objects.

Live Markers

Live Markers is another hybrid approach that takes advantage from the superior ability of LWOF on weakly defined segments and from IFTSC to handle complex 2D/3D shapes of multiple objects.

The markers may be selected by the user or may come from the live-wire segments.

Live Markers

In several cases, the live-wire segments followed by IFTSC almost complete the segmentation process.

Live Markers

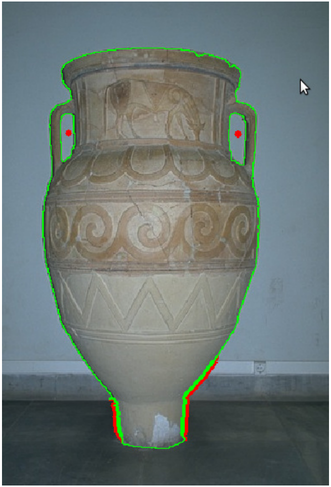
In several cases, the live-wire segments followed by IFTSC almost complete the segmentation process.

Live Markers

In several cases, the live-wire segments followed by IFTSC almost complete the segmentation process.

Live Markers

In several cases, the live-wire segments followed by IFTSC almost complete the segmentation process.



Organization of this lecture

- Object delineation using the image foresting transform.
- A comparison between IFTSC and object delineation using the min-cut/max-flow algorithm (GCMF).
- Fuzzy object models and video segmentation.

IFTSC and min-cut/max-flow algorithms

- It can be shown that the IFTSC computes the graph cut whose minimum arc weight

$$\min_{\forall (s,t) \in \mathcal{A}, L(s)=1, L(t)=0} w(s,t)$$

is maximum, considering all possible cuts between internal and external seeds, and this is also a piecewise optimum cut.

IFTSC and min-cut/max-flow algorithms

- It can be shown that the IFTSC computes the graph cut whose minimum arc weight

$$\min_{\forall(s,t)\in\mathcal{A},L(s)=1,L(t)=0} w(s,t)$$

is maximum, considering all possible cuts between internal and external seeds, and this is also a piecewise optimum cut.

- Similar region-based delineation could be obtained by GCMF as a graph cut whose sum of arc weights

$$\sqrt[\beta]{\sum_{\forall(s,t)\in\mathcal{A}|\ L(s)=1,L(t)=0} \bar{w}^{\beta}(s,t)}$$

is minimum for $\beta \geq 1$, with lower values favoring smaller cuts and higher values making both equivalent.

IFTSC and min-cut/max-flow algorithms

- IFTSC can handle multiple object delineation in $O(|\mathcal{N}|)$.

IFTSC and min-cut/max-flow algorithms

- IFTSC can handle multiple object delineation in $O(|\mathcal{N}|)$.
- GCMF is not viable for multiple objects and takes $O(|\mathcal{N}|^{2.5})$ for single object delineation.

IFTSC and min-cut/max-flow algorithms

- IFTSC can handle multiple object delineation in $O(|\mathcal{N}|)$.
- GCMF is not viable for multiple objects and takes $O(|\mathcal{N}|^{2.5})$ for single object delineation.
- IFTSC is also more robust with respect to seed location than GCMF, but the latter provides smoother boundaries with less leaking than the former.

IFTSC and min-cut/max-flow algorithms

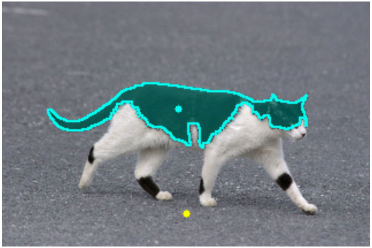
- IFTSC can handle multiple object delineation in $O(|\mathcal{N}|)$.
- GCMF is not viable for multiple objects and takes $O(|\mathcal{N}|^{2.5})$ for single object delineation.
- IFTSC is also more robust with respect to seed location than GCMF, but the latter provides smoother boundaries with less leaking than the former.
- Interestingly, GCMF and LWOF are known to be related by dual graphs just as IFTSC and Riverbed.

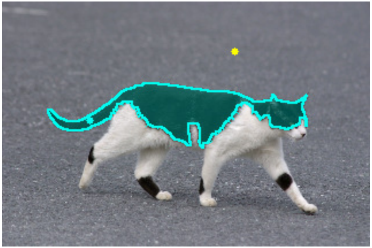
IFTSC and min-cut/max-flow algorithms

A lower β value allows GCMF (left) to obtain a smoother boundary, reducing the leaking of IFTSC (right).

IFTSC and min-cut/max-flow algorithms

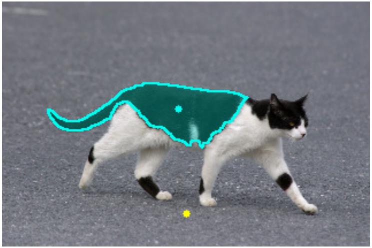
However, a same connected component is always obtained with IFTSC, independently of seed location.

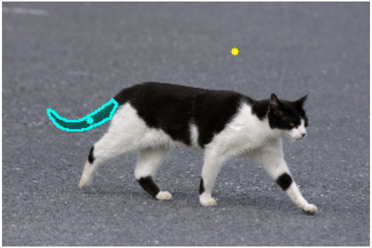




IFTSC and min-cut/max-flow algorithms

The same does not happen in GCMF, when β is not high enough.





Graph Cuts using the Min-Cut/Max-Flow algorithm

Some GCMF approaches further change the graph topology by adding two virtual *source* and *sink* nodes, connected to every node in the graph

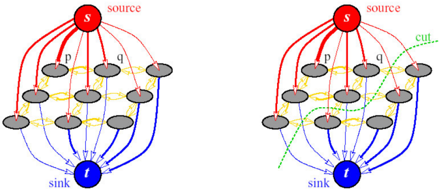


Image source: mathworks.com

and consider the following energy function in order to circumvent the drawbacks of smaller cuts

$$\sum_{\forall (s,t)\in\mathcal{A}|\ L(s)=1, L(t)=0} \bar{w}(s,t) + \lambda \left(\sum_{\forall (s)\in\mathcal{I}|L(s)=1} \bar{P}_o(s) + \sum_{\forall (t)\in\mathcal{I}|L(t)=0} \bar{P}_b(t) \right),$$

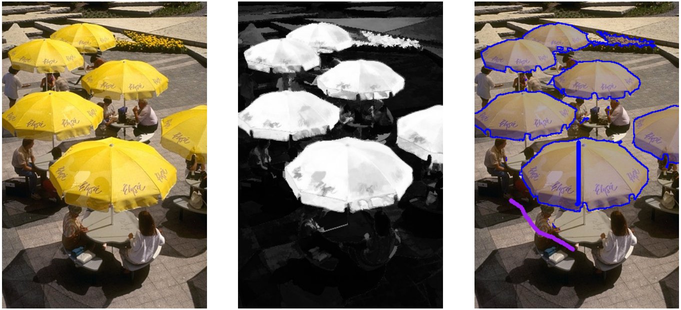
where P is an object membership map (probability map).

Graph Cuts using the Min-Cut/Max-Flow algorithm

- However, this leads to a dependence on the quality of the object membership map P.

Graph Cuts using the Min-Cut/Max-Flow algorithm

- However, this leads to a dependence on the quality of the object membership map P.
- Moreover, the addition of the virtual nodes makes it difficult to guarantee that the resulting segmentation will be spatially connected to the user-drawn markers.



Graph Cuts using the Min-Cut/Max-Flow algorithm

- However, this leads to a dependence on the quality of the object membership map P.
- Moreover, the addition of the virtual nodes makes it difficult to guarantee that the resulting segmentation will be spatially connected to the user-drawn markers.

- Lower λ values imply the regular GCMF minimum cut measure while higher values lead to simple thresholding of P.

Organization of this lecture

- Object delineation using the image foresting transform.
- A comparison between IFTSC and object delineation using the min-cut/max-flow algorithm (GCMF).
- Fuzzy object models and video segmentation.

Statistical object models

- Automatic segmentation is feasible when possible object deformations are captured into a statistical model (atlas).

Statistical object models

- Automatic segmentation is feasible when possible object deformations are captured into a statistical model (atlas).
- The atlas is built by registration among training images in order to estimate the probability of each spel to be inside object/background.

Statistical object models

- Automatic segmentation is feasible when possible object deformations are captured into a statistical model (atlas).
- The atlas is built by registration among training images in order to estimate the probability of each spel to be inside object/background.
- Object location in a test image is solved when it is registered with the atlas.

Statistical object models

- Automatic segmentation is feasible when possible object deformations are captured into a statistical model (atlas).
- The atlas is built by registration among training images in order to estimate the probability of each spel to be inside object/background.
- Object location in a test image is solved when it is registered with the atlas.
- Subsequent object delineation completes segmentation.

Statistical object models

- Automatic segmentation is feasible when possible object deformations are captured into a statistical model (atlas).
- The atlas is built by registration among training images in order to estimate the probability of each spel to be inside object/background.
- Object location in a test image is solved when it is registered with the atlas.
- Subsequent object delineation completes segmentation.

Registration is an expensive task that may force delineation to fit with the model irrespective to the local image information.

Fuzzy object models

We have developed fuzzy models to eliminate registration and provide more decision power to the delineation method.

Fuzzy object models

We have developed fuzzy models to eliminate registration and provide more decision power to the delineation method.

- A fuzzy model may only require a simple translation of the training objects to a common reference point for its construction.

Fuzzy object models

We have developed fuzzy models to eliminate registration and provide more decision power to the delineation method.

- A fuzzy model may only require a simple translation of the training objects to a common reference point for its construction.
- It may also require object alignment, but this only involves its own image.

Fuzzy object models

We have developed fuzzy models to eliminate registration and provide more decision power to the delineation method.

- A fuzzy model may only require a simple translation of the training objects to a common reference point for its construction.
- It may also require object alignment, but this only involves its own image.
- Segmentation is solved by translating the model and executing delineation at each location.

Fuzzy object models

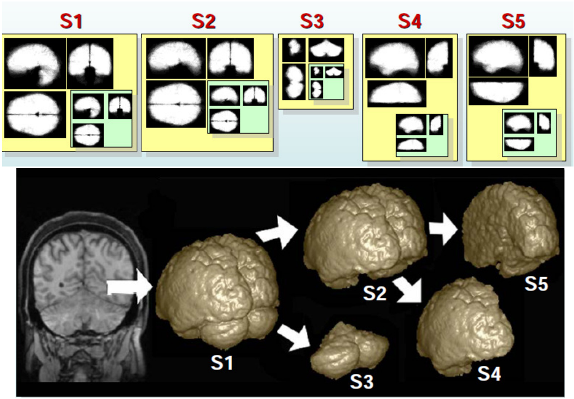
We have developed fuzzy models to eliminate registration and provide more decision power to the delineation method.

- A fuzzy model may only require a simple translation of the training objects to a common reference point for its construction.
- It may also require object alignment, but this only involves its own image.
- Segmentation is solved by translating the model and executing delineation at each location.
- It is possible to considerably speed-up this object search process by using multiple scales of the image and models.

Fuzzy object models

Examples of objects and their fuzzy models.

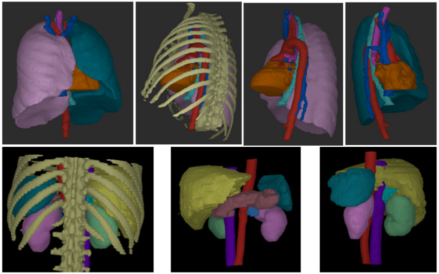
Medical imaging: Object modeling and image segmentation



This lecture will present only the first case, named Cloud System Model (CSM), using IFTSC for delineation.

Fuzzy object models

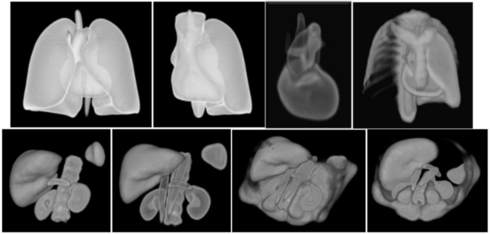
Examples of objects and their fuzzy models.



This lecture will present only the first case, named Cloud System Model (CSM), using IFTSC for delineation.

Fuzzy object models

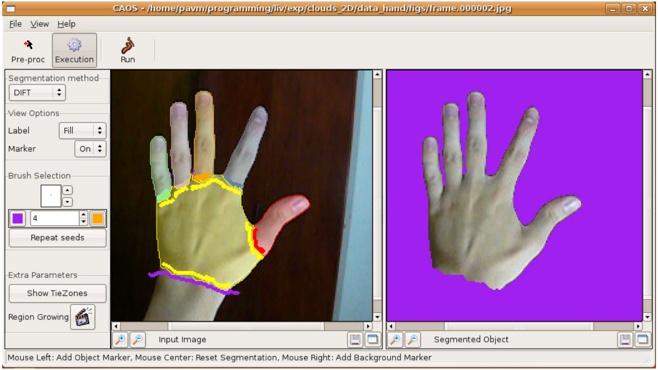
Examples of objects and their fuzzy models.



This lecture will present only the first case, named Cloud System Model (CSM), using IFTSC for delineation.

The cloud system model

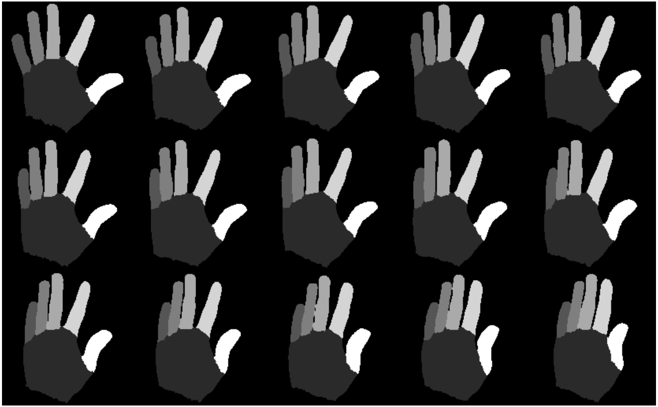
A set of training objects is first provided by interactive IFTSC segmentation.



Each image with multiple objects forms an object system with a common reference point (e.g., the geometric center of the objects).

The cloud system model

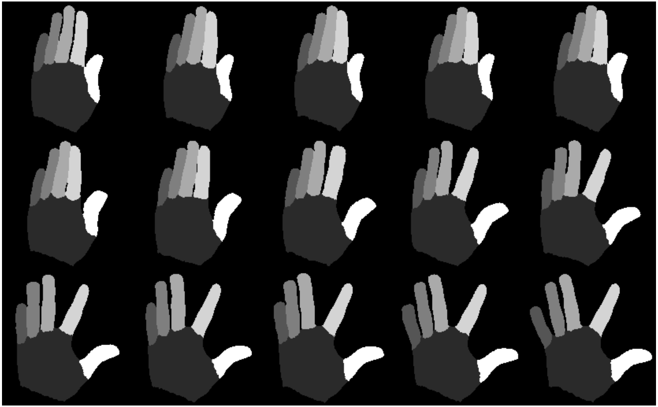
A set of training objects is first provided by interactive IFTSC segmentation.



Each image with multiple objects forms an object system with a common reference point (e.g., the geometric center of the objects).

The cloud system model

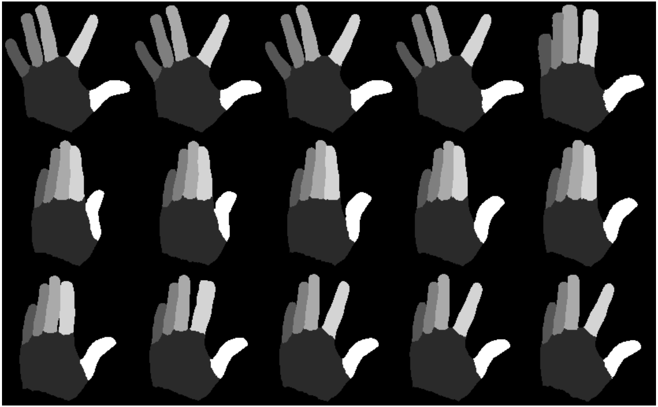
A set of training objects is first provided by interactive IFTSC segmentation.



Each image with multiple objects forms an object system with a common reference point (e.g., the geometric center of the objects).

The cloud system model

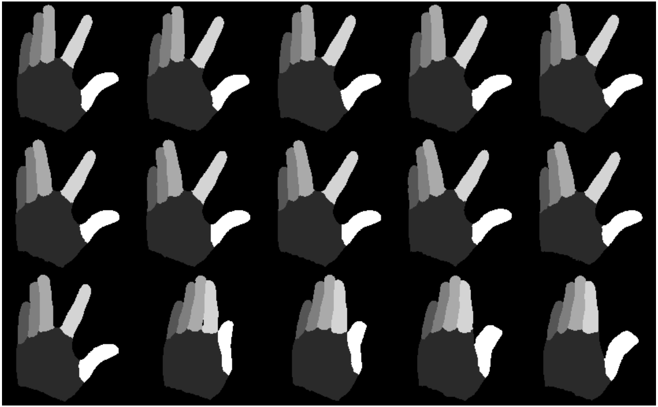
A set of training objects is first provided by interactive IFTSC segmentation.



Each image with multiple objects forms an object system with a common reference point (e.g., the geometric center of the objects).

The cloud system model

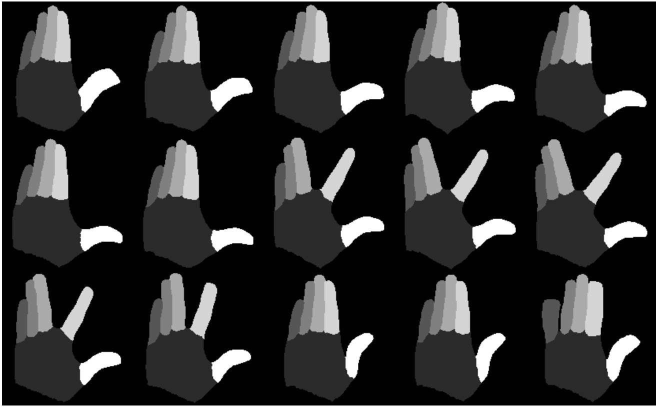
A set of training objects is first provided by interactive IFTSC segmentation.



Each image with multiple objects forms an object system with a common reference point (e.g., the geometric center of the objects).

The cloud system model

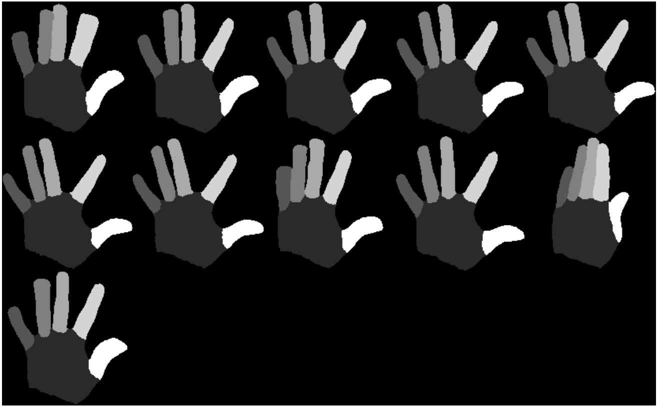
A set of training objects is first provided by interactive IFTSC segmentation.



Each image with multiple objects forms an object system with a common reference point (e.g., the geometric center of the objects).

The cloud system model

A set of training objects is first provided by interactive IFTSC segmentation.



Each image with multiple objects forms an object system with a common reference point (e.g., the geometric center of the objects).

The cloud system model

Groups of object systems in which the corresponding objects have similar shapes, sizes and positions form different cloud system models, as follows.

The cloud system model

Groups of object systems in which the corresponding objects have similar shapes, sizes and positions form different cloud system models, as follows.

- Each object system becomes a node of a complete graph, where the weight of each arc derives from the similarities between the corresponding objects in shape, size and position.

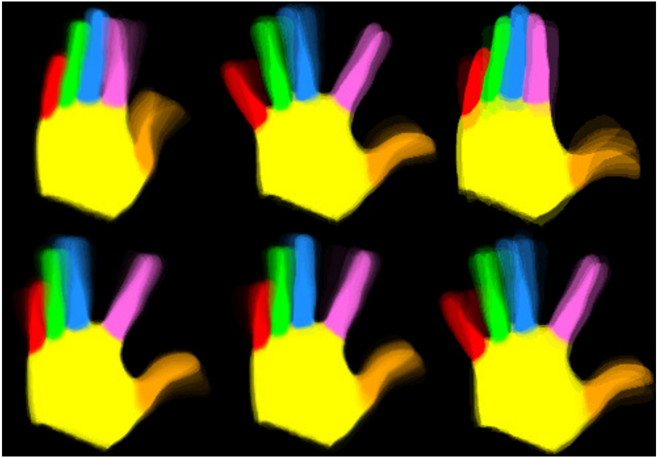
The cloud system model

Groups of object systems in which the corresponding objects have similar shapes, sizes and positions form different cloud system models, as follows.

- Each object system becomes a node of a complete graph, where the weight of each arc derives from the similarities between the corresponding objects in shape, size and position.
- The groups are found as maximal cliques in which all arc weights are higher than a threshold.

The cloud system model

The object systems in each group are finally translated to a same reference point and the corresponding object masks are averaged, forming a set of cloud systems.



The cloud system model

A cloud system model (CSM) then consists of three elements:

The cloud system model

A cloud system model (CSM) then consists of three elements:

- A fuzzy membership map (object clouds), which indicates an object uncertainty region with values strictly lower than 1 and higher than 0.

The cloud system model

A cloud system model (CSM) then consists of three elements:

- A fuzzy membership map (object clouds), which indicates an object uncertainty region with values strictly lower than 1 and higher than 0.
- A delineation algorithm (IFTSC), whose execution is constrained in the uncertainty region.

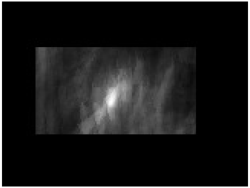
The cloud system model

A cloud system model (CSM) then consists of three elements:

- A fuzzy membership map (object clouds), which indicates an object uncertainty region with values strictly lower than 1 and higher than 0.
- A delineation algorithm (IFTSC), whose execution is constrained in the uncertainty region.
- A criterion function, which assigns a score to any set of delineated objects.

The cloud system model

Segmentation using CSM consists of a search for the translation to the image location which produces the highest score, when the reference point of the most suitable cloud system is at that position.



show video handsearch.mpg

The cloud system model in video

Video object segmentation is a challenging problem due to

The cloud system model in video

Video object segmentation is a challenging problem due to

- The presence of intra- and inter-object occlusions (partial or total),

The cloud system model in video

Video object segmentation is a challenging problem due to

- The presence of intra- and inter-object occlusions (partial or total),
- Deformable and articulated objects,

The cloud system model in video

Video object segmentation is a challenging problem due to

- The presence of intra- and inter-object occlusions (partial or total),
- Deformable and articulated objects,
- Motion,

The cloud system model in video

Video object segmentation is a challenging problem due to

- The presence of intra- and inter-object occlusions (partial or total),
- Deformable and articulated objects,
- Motion,
- Color and texture similarities with the background,

The cloud system model in video

Video object segmentation is a challenging problem due to

- The presence of intra- and inter-object occlusions (partial or total),
- Deformable and articulated objects,
- Motion,
- Color and texture similarities with the background,
- Poor illumination.

The cloud system model in video

- The CSM extension to video was developed for aiding the diagnosis of toddlers at risk of autism.

The cloud system model in video

- The CSM extension to video was developed for aiding the diagnosis of toddlers at risk of autism.
- Hence, we have first tackled the problem of dealing with articulated objects (i.e., the human body).

The cloud system model in video

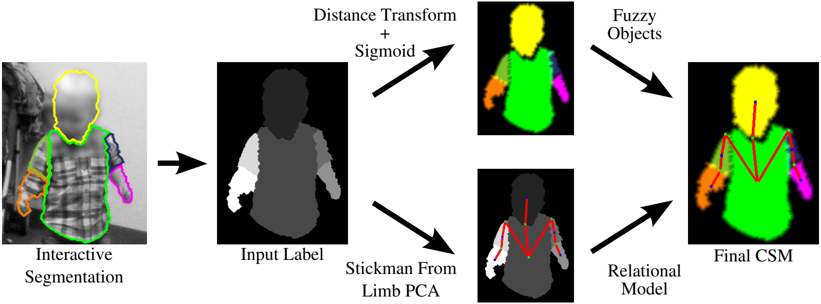
- The CSM extension to video was developed for aiding the diagnosis of toddlers at risk of autism.
- Hence, we have first tackled the problem of dealing with articulated objects (i.e., the human body).
- Given the wide range of possible body poses, we start from a single segmentation mask obtained interactively in the first frame.

The cloud system model in video

- The CSM extension to video was developed for aiding the diagnosis of toddlers at risk of autism.
- Hence, we have first tackled the problem of dealing with articulated objects (i.e., the human body).
- Given the wide range of possible body poses, we start from a single segmentation mask obtained interactively in the first frame.
- The remainder of the video is automatically segmented using CSM.

The cloud system model in video

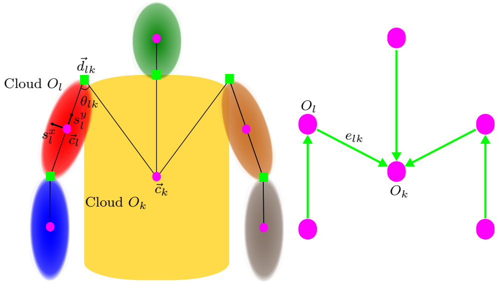
Creation of a CSM from a single segmentation mask in the first frame.



This process outputs a CSM with one cloud per body part.

The cloud system model in video

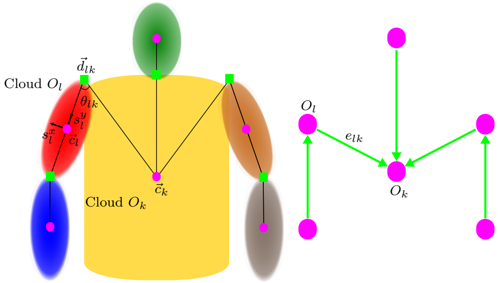
To deal with articulation, we have coupled the CSM with a relational model that is used to reposition the clouds during the search in a new test frame.



- The torso translates over the image and carries along the head and limbs.
- While the torso can be globally rotated, the remaining body part joint angles are relative between the connected parts.
- The scale is relative to the first frame for all body parts.

The cloud system model in video

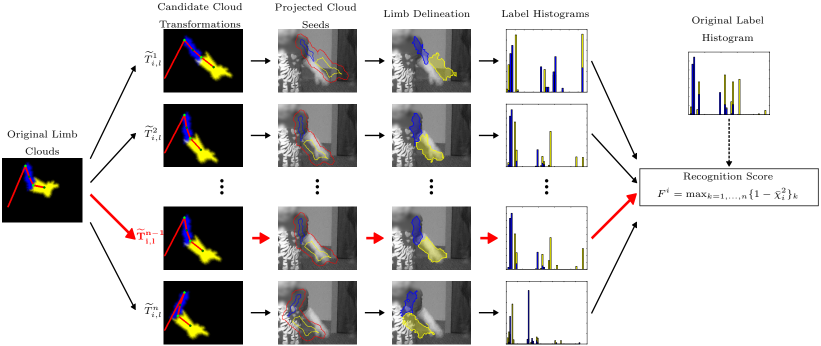
To deal with articulation, we have coupled the CSM with a relational model that is used to reposition the clouds during the search in a new test frame.



- The torso translates over the image and carries along the head and limbs.
- While the torso can be globally rotated, the remaining body part joint angles are relative between the connected parts.
- The scale is relative to the first frame for all body parts.
- The search for the body in a new frame consists of finding an optimal configuration for those parameters using multi-scale parameter search (MSPS).

The cloud system model in video

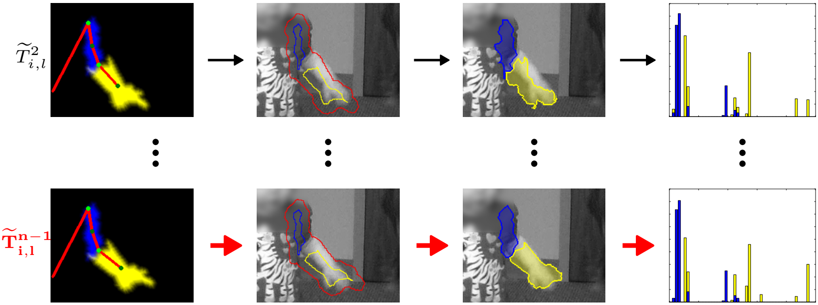
For every configuration of the relational model parameters chosen by MSPS, the CSM performs delineation and evaluates the criterion function.



In this case, the criterion function considers the χ^2 distance between color histograms extracted from both the first and the current frame.

The cloud system model in video

Delineation uses the regular IFTSC with the pixels surrounding the uncertainty region as seeds.



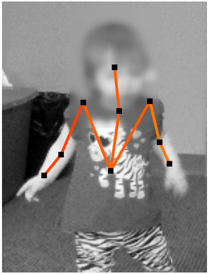
But how does this aid autism assessment?

Autism Spectrum Disorder (ASD) is a behavioral disorder whose symptoms range from lack of eye contact to abnormal motor behavior (e.g., body rocking, arm-and-hand flapping).

But how does this aid autism assessment?

Autism Spectrum Disorder (ASD) is a behavioral disorder whose symptoms range from lack of eye contact to abnormal motor behavior (e.g., body rocking, arm-and-hand flapping).

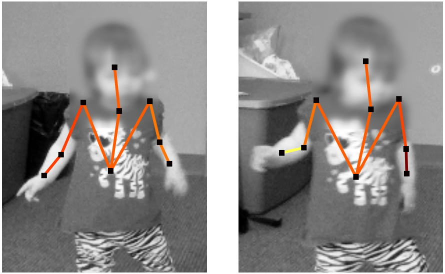
- The human segmentation using CSM naturally provides us 2D body pose estimation from the relational model used to guide it.



But how does this aid autism assessment?

Autism Spectrum Disorder (ASD) is a behavioral disorder whose symptoms range from lack of eye contact to abnormal motor behavior (e.g., body rocking, arm-and-hand flapping).

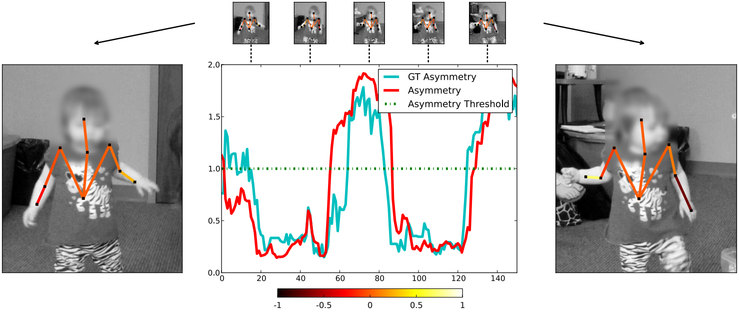
- The human segmentation using CSM naturally provides us 2D body pose estimation from the relational model used to guide it.



- We use the 2D body pose to detect arm asymmetry during unsupported gait, a possible sign of autism in toddlers, from videos of real in-clinic ASD assessments.

But how does this aid autism assessment?

We provide quantifiable measures to aid the clinician in his/her assessment, which can be used for both research and diagnosis.



Conclusion

- It should be clear the importance of combining the strengths from distinct object delineation methods.

Conclusion

- It should be clear the importance of combining the strengths from distinct object delineation methods.
- The synergistic combination between object models and delineation methods makes automatic segmentation feasible.

Conclusion

- It should be clear the importance of combining the strengths from distinct object delineation methods.
- The synergistic combination between object models and delineation methods makes automatic segmentation feasible.
- When any automatic segmentation method fails, we have also developed solutions to reduce it into an optimum-path forest with minimum number of roots, so this facilitates corrections by the differential IFTSC algorithm.