Tempo-spatial Segmentation of Mesh Sequence*

* Submission to Computers & Graphics (Elsevier), Nov. 2012

PhD student: Guoliang LUO

gluo@unistra.fr

December 6, 2012

UNIVERSITÉ DE STRASBOURG

Outline

Motivation

'Facial-expression' data obtained from motion capture:

Applications

- Compression.
- Shape query.
- Motion recognition.
- ...

Segmentation on the mesh sequence provides a solution to the above applications.

Related works

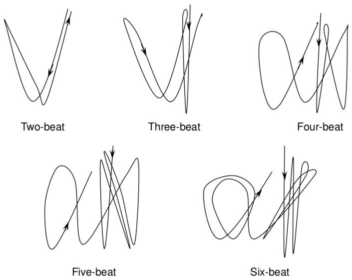
Temporal segmentation / Spatial segmentation

On video

- [Wang et al.]'01

On pose data

- [Fod et al.]'02
- [Barbic et al.]'04



[Wang et al.]'01

Related works

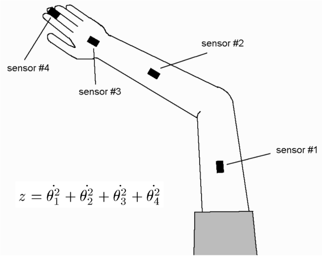
Temporal segmentation / Spatial segmentation

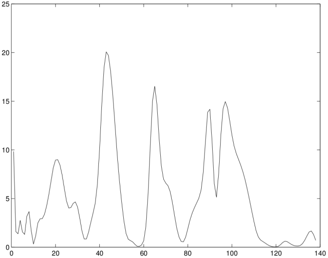
On video

- [Wang et al.]'01

On pose data

- [Fod et al.]'02
- [Barbic et al.]'04





Sum of squares of angular velocity. [Fod et al.]'02

Related works

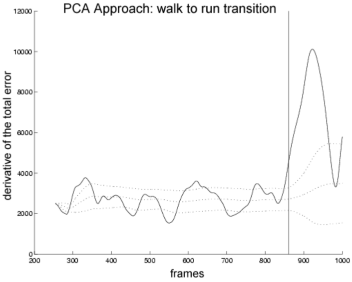
Temporal segmentation / Spatial segmentation

On video

- [Wang et al.]'01

On pose data

- [Fod et al.]'02
- [Barbic et al.]'04



Reconstruction errors. [Barbic et al.]'04

Related works

Temporal segmentation / Spatial segmentation

Spectral clustering

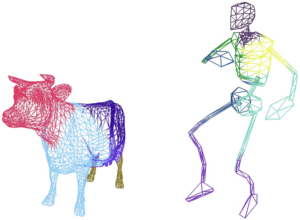
- [Sattler et al.]'05
- [Kalafatlar et al.]'10
- *Entire sequence*

Region growing

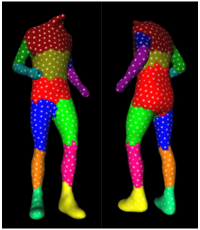
- [Lee et al.]'06
- Per-frame based

Learning

- [Kalogerakis et al.]'10
- [Benhabiles et al.]'11
- Ground-truth sets
- Not dynamic mesh



(c) [Sattler et al.]'05



(d) [Kalafatlar et al.]'10

Related works

Temporal segmentation / Spatial segmentation

Spectral clustering

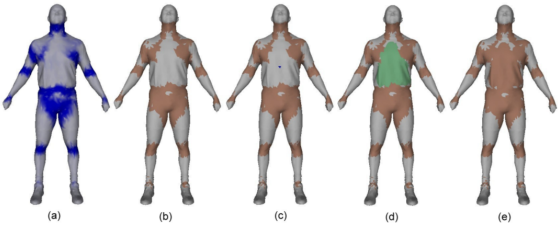
- [Sattler et al.]'05
- [Kalafatlar et al.]'10
- Entire sequence

Region growing

- [Lee et al.]'06
- *Per-frame based*

Learning

- [Kalogerakis et al.]'10
- [Benhabiles et al.]'11
- Ground-truth sets
- Not dynamic mesh



[Lee et al.]'06

Related works

Temporal segmentation / Spatial segmentation

Spectral clustering

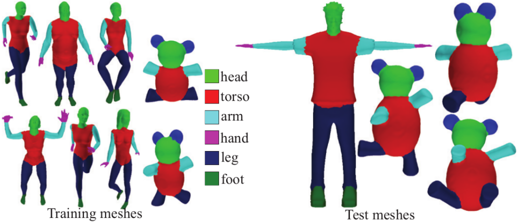
- [Sattler et al.]'05
- [Kalafatlar et al.]'10
- Entire sequence

Region growing

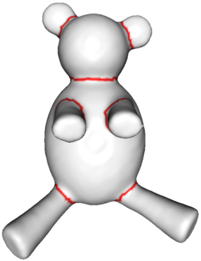
- [Lee et al.]'06
- Per-frame based

Learning

- [Kalogerakis et al.]'10
- [Benhabiles et al.]'11
- *Ground-truth sets*
- *Not dynamic mesh*



(e) [Kalogerakis et al.]'10

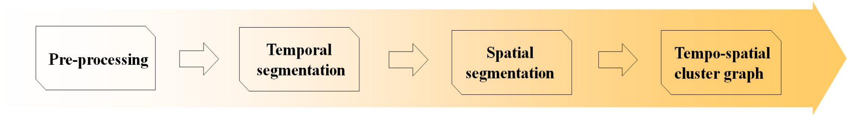


(f) [Benhabiles et al.]'11

Objectives

- Temporal segmentation of **mesh sequence**.
- Spatial segmentation within each temporal segment.
- High-level representation of mesh sequence.

Tempo-spatial segmentation of mesh sequence



Overview of our approach.

Temporal segmentation

- Maximizing within-segment frame affinities.

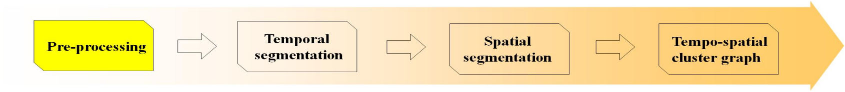
Spatial segmentation

- Deformation based.

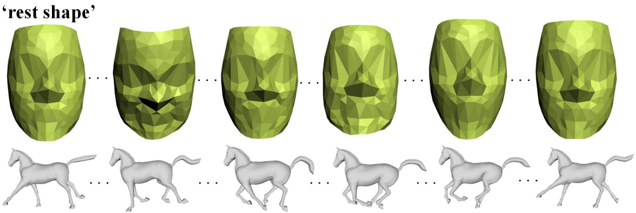
Tempo-spatial cluster graph

- Devising a high-level representation of mesh sequence.

Pre-processing

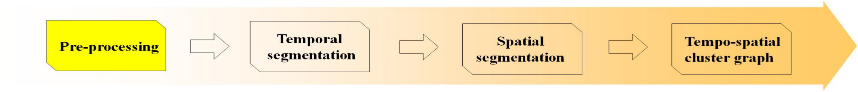


- Input data

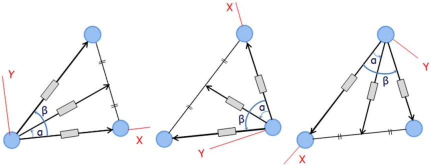


- Deformation feature descriptor: Triangle-based strain
- Geodesic distance:
  - Triangle pair
  - Averaging 9-pair of vertex-distance

Pre-processing



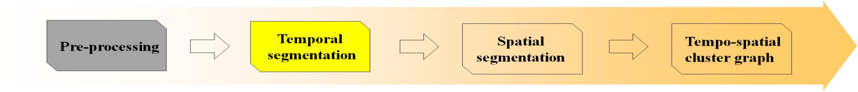
- Input data
- Deformation feature descriptor: Triangle-based strain



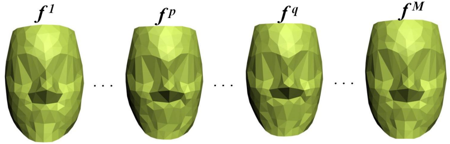
Triangle-based strain approximated by using linear gage rosettes.[Seo et. al]'12

- Geodesic distance:
  - Triangle pair
  - Averaging 9-pair of vertex-distance

Frame affinity



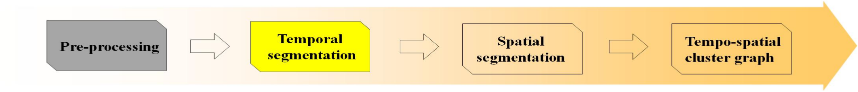
Frame affinity $A_f(p,q)$, $p,q = 1,\ldots,M$:



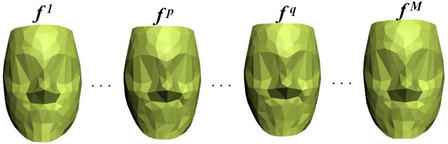
$$\mathbf{s^p} = (s_1^p, ..., s_N^p)^T, \ \mathbf{s^q} = (s_1^q, ..., s_N^q)^T$$
$$A_f(p,q) = exp(-0.5\delta_t^2 \parallel \mathbf{s^p} - \mathbf{s^q} \parallel_{L_2}^2)$$

- M:Number of frames,
- N:Number of triangles,
- s_i^p:The strain of the i-th triangle in the p-th frame.

Frame affinity



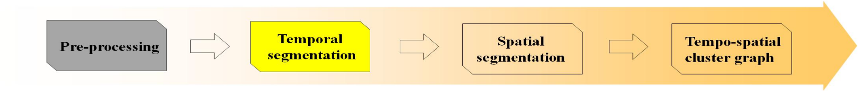
Frame affinity $A_f(p,q)$, $p,q = 1,\ldots,M$:



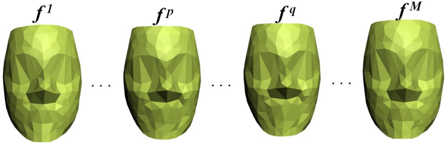
$$\mathbf{s^p} = (s_1^p, \ldots, s_N^p)^T,\ \mathbf{s^q} = (s_1^q, \ldots, s_N^q)^T$$
$$A_f(p,q) = exp(-0.5\delta_t^2 \parallel \mathbf{s^p} - \mathbf{s^q} \parallel_{L_2}^2)$$

- M:Number of frames,
- N:Number of triangles,
- s_i^p:The strain of the i-th triangle in the p-th frame.

Frame affinity



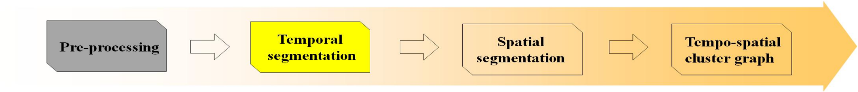
Frame affinity $A_f(p,q)$, $p,q = 1,\ldots,M$:



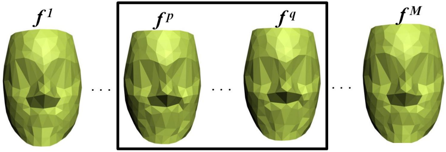
$$\mathbf{s^p} = (s_1^p, \ldots, s_N^p)^T, \ \mathbf{s^q} = (s_1^q, \ldots, s_N^q)^T$$
$$A_f(p,q) = exp(-0.5\delta_t^2 \parallel \mathbf{s^p} - \mathbf{s^q} \parallel_{L_2}^2)$$

•M:Number of frames,
•N:Number of triangles,
•s_i^p:The strain of the i-th triangle in the p-th frame.

Average frame affinity

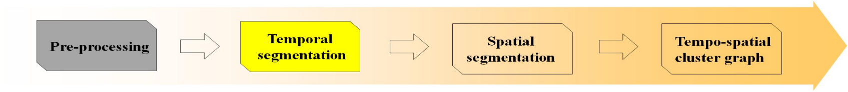


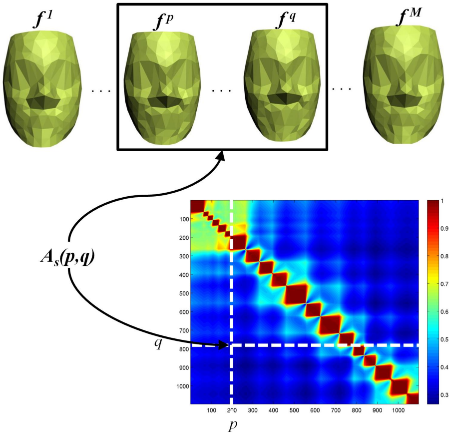
Average frame affinity $A_s(p,q)$ within subsequence $[p,q]$:



$$A_s(p,q) = \frac{\sum_{m=n+1}^{q} \sum_{n=p}^{q-1} A_f(n,m)}{(q-p+1)\cdot(q-p)/2}, p < q.$$

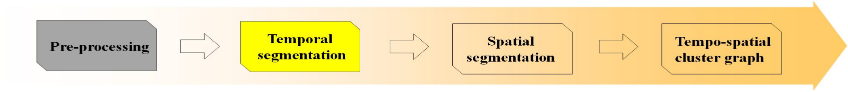
Average frame-affinity matrix

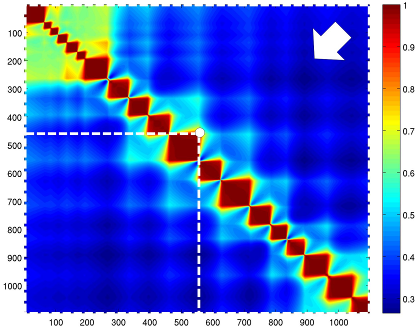




An example of $A_s(p, q)$ in an average frame-affinity matrix.

Temporal segmentation





(g) *Average frame-affinity matrix.*

Algorithm 1: $TempSeg(I_B, A_s, I_h, I_t)$

Init: $I_B=[]$, $I_h=1$, $I_t=M$, A_s

$L=I_t-I_h+1$

for $l=L$ to 1 **do**

  for $p=1$ to $L-l+1$ **do**

    $A_{s\text{-}sub}(p)=A_s(p,p+l-1)$

  end for

  $[A_{s\text{-}max}, p]=\mathbf{max}(A_{s\text{-}sub})$

  if $A_{s\text{-}max}>\rho_0$ **then**

    $I_B=[I_B\ p]$

    $q=p+l-1;$

    $TempSeg(I_B, A_s, I_h, p-1)$

    $TempSeg(I_B, A_s, q+1, I_t)$

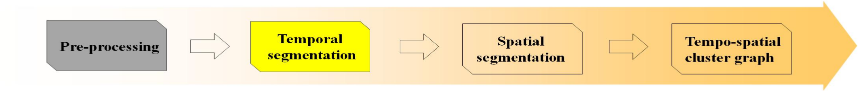
    Break

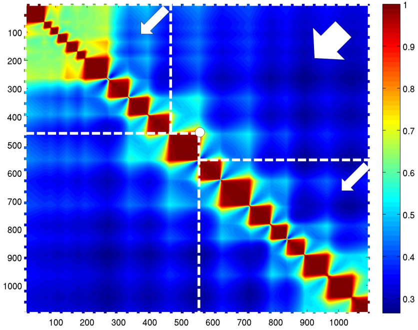
  end if

end for

Return: I_B

Temporal segmentation





(i) *Average frame-affinity matrix.*

Algorithm 1: $TempSeg(I_B, A_s, I_h, I_t)$

Init: I_B=[], I_h=1, I_t=M, A_s

$L=I_t-I_h+1$

for $l=L$ to 1 **do**

  for p=1 to $L-l+1$ **do**

    $A_{s\text{-}sub}(p)=A_s(p,p+l-1)$

  end for

  $[A_{s\text{-}max}, p]$=**max**$(A_{s\text{-}sub})$

  if $A_{s\text{-}max}>\rho_0$ **then**

    $I_B=[I_B\ p]$

    $q=p+l-1;$

    $TempSeg(I_B,A_s,I_h,p-1)$

    $TempSeg(I_B,A_s,q+1,I_t)$

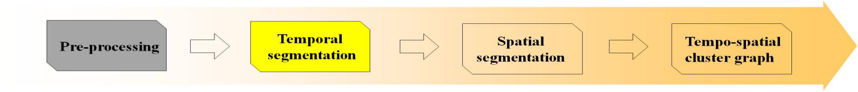
    Break

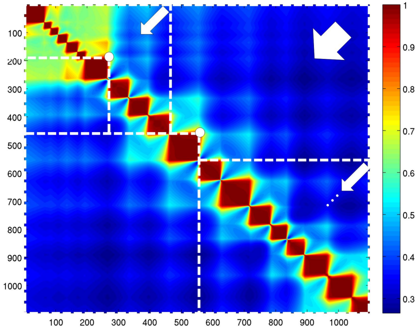
  end if

end for

Return: I_B

Temporal segmentation



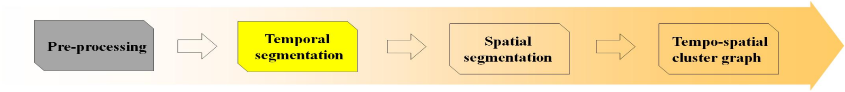


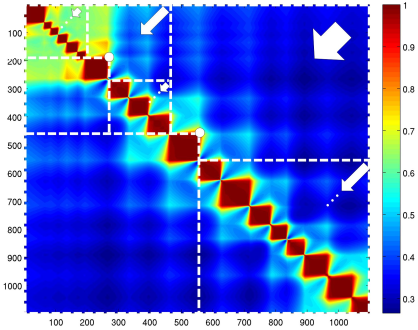
(k) *Average frame-affinity matrix.*

Algorithm 1: $TempSeg(I_B, A_s, I_h, I_t)$

Init: I_B=[], I_h=1, I_t=M, A_s
    $L=I_t-I_h+1$
    for $l=L$ to 1 **do**
        for p=1 to $L-l+1$ **do**
            $A_{s\text{-}sub}(p)=A_s(p,p+l-1)$
        end for
        $[A_{s\text{-}max}, p]$=**max**$(A_{s\text{-}sub})$
        if $A_{s\text{-}max}>\rho_0$ **then**
            $I_B=[I_B\ p]$
            $q=p+l-1;$
            $TempSeg(I_B, A_s, I_h, p-1)$
            $TempSeg(I_B, A_s, q+1, I_t)$
            Break
        end if
    end for
    Return: I_B

Temporal segmentation



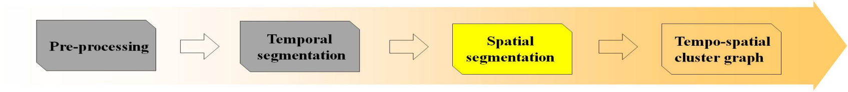


(m) *Average frame-affinity matrix.*

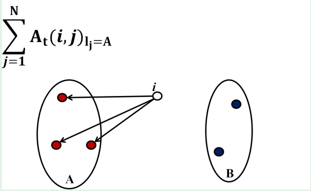
Algorithm 1: $TempSeg(I_B, A_s, I_h, I_t)$

Init: I_B=[], I_h=1, I_t=M, A_s
    $L=I_t-I_h+1$
    for $l=L$ to 1 **do**
        for $p=1$ to $L-l+1$ **do**
            $A_{s\text{-}sub}(p)=A_s(p,p+l-1)$
        end for
        $[A_{s\text{-}max}, p]$=**max**$(A_{s\text{-}sub})$
        if $A_{s\text{-}max}>\rho_0$ **then**
            $I_B=[I_B\ p]$
            $q=p+l-1;$
            $TempSeg(I_B, A_s, I_h, p-1)$
            $TempSeg(I_B, A_s, q+1, I_t)$
            Break
        end if
    end for
    Return: I_B

Initial segmentation



Correlation clustering ([Bagon et. al]'11)



Triangle affinities

$\forall i, j = 1, \ldots, N$,

$$dist_G(i,j) \xrightarrow{GKF} A'_t(i,j) \xrightarrow{\substack{linearly \\ rescale}} A_t(i,j),$$

where $A'_t(i,j) \in [0\ 1]$ and $A_t(i,j) \in [-1\ 1]$.

Algorithm 2: *Adaptive-label ICM*

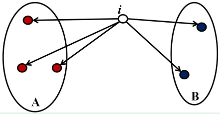
Init: l_i=1, i=1,..., n, L=1, A_t
- **repeat**
  - **for** i=1 to n **do**
    - **for** l=1 to L **do**
      - $$A_l = \sum_{j=1}^{n} A_t(i,j)_{l_j=l}$$
    - **end for**
    - **if** $\forall A_l < 0$ **then**
      - L=L+1
      - l_i=L
    - **else**
      - [$A_{l\text{-}max}$,l_{max}]=**max**(A_l)
      - l_i=l_{max}
    - **end if**
  - **end for**
- **until** L is unchanged
- **return** l_i, i=1,…, n

Initial segmentation

Pre-processing → Temporal segmentation → **Spatial segmentation** → Tempo-spatial cluster graph

Correlation clustering ([Bagon et. al]'11)

$$\sum_{j=1}^{N} \mathbf{A_t}(\boldsymbol{i},\boldsymbol{j})_{\mathbf{l_j=A}} \qquad \sum_{j=1}^{N} \mathbf{A_t}(\boldsymbol{i},\boldsymbol{j})_{\mathbf{l_j=B}}$$



Triangle affinities

$\forall i, j = 1, \ldots, N,$

$$dist_G(i,j) \xrightarrow{GKF} A_t'(i,j) \xrightarrow[]{\substack{linearly\\rescale}} A_t(i,j),$$

where $A_t'(i,j) \in [0\ 1]$ and $A_t(i,j) \in [-1\ 1]$.

Algorithm 2: *Adaptive-label ICM*

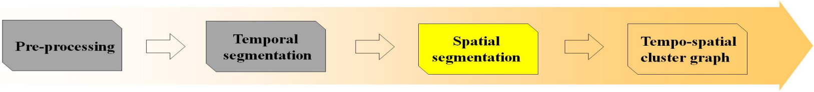
Init: l_i=1, i=1,..., n, L=1, A_t

repeat
- **for** i=1 to n **do**
  - **for** l=1 to L **do**
    - $A_l = \sum_{j=1}^{n} A_t(i,j)_{l_j=l}$
  - **end for**
  - **if** $\forall A_l < 0$ **then**
    - L=L+1
    - l_i=L
  - **else**
    - $[A_{l\text{-}max}, l_{max}]$=**max**$(A_l)$
    - l_i=l_{max}
  - **end if**
- **end for**

until L is unchanged

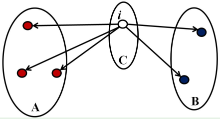
return l_i, i=1,..., n

Initial segmentation



Correlation clustering ([Bagon et. al]'11)

$$\sum_{j=1}^{N} \mathbf{A_t(i,j)_{l_j=A}} < \mathbf{0} \quad \sum_{j=1}^{N} \mathbf{A_t(i,j)_{l_j=B}} < \mathbf{0}$$



Triangle affinities

$\forall i, j = 1, \ldots, N,$

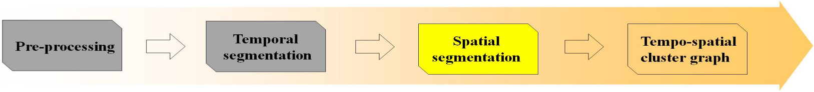
$$dist_G(i,j) \xrightarrow{GKF} A'_t(i,j) \xrightarrow[\text{rescale}]{\text{linearly}} A_t(i,j),$$

where $A'_t(i,j) \in [0\ 1]$ and $A_t(i,j) \in [-1\ 1]$.

Algorithm 2: *Adaptive-label ICM*

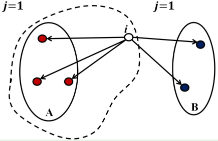
Init: l_i=1, i=1,..., n, L=1, A_t
  repeat
    for i=1 to n **do**
      for l=1 to L **do**
        $A_l = \sum_{j=1}^{n} A_t(i,j)_{l_j=l}$
      end for
      if $\forall A_l < 0$ **then**
        L=L+1
        l_i=L
      else
        [$A_{l\text{-}max}$,l_{max}]=**max**(A_l)
        l_i=l_{max}
      end if
    end for
  until L is unchanged
  return l_i, i=1,..., n

Initial segmentation



Correlation clustering ([Bagon et. al]'11)

$$\sum_{j=1}^{N} \mathbf{A_t}(\boldsymbol{i},\boldsymbol{j})_{\mathbf{l_j=A}} > \sum_{j=1}^{N} \mathbf{A_t}(\boldsymbol{i},\boldsymbol{j})_{\mathbf{l_j=B}}$$



Triangle affinities

$\forall i, j = 1, \ldots, N,$

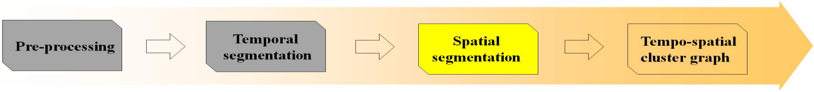
$$dist_G(i,j) \xrightarrow{GKF} A'_t(i,j) \xrightarrow[\text{}]{\substack{linearly\\ rescale}} A_t(i,j),$$

where $A'_t(i,j) \in [0\ 1]$ and $A_t(i,j) \in [-1\ 1]$.

Algorithm 2: *Adaptive-label ICM*

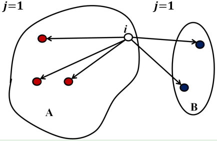
Init: l_i=1, i=1,..., n, L=1, A_t
  repeat
    for i=1 to n **do**
      for l=1 to L **do**
$$A_l = \sum_{j=1}^{n} A_t(i,j)_{l_j=l}$$
      end for
      if $\forall A_l < 0$ **then**
        L=L+1
        l_i=L
      else
        $[A_{l\text{-}max}, l_{max}]$=**max**$(A_l)$
        l_i=l_{max}
      end if
    end for
  until L is unchanged
  return l_i, i=1,..., n

Initial segmentation



Correlation clustering ([Bagon et. al]'11)

$$\sum_{j=1}^{N} A_t(i,j)_{l_j=A} > \sum_{j=1}^{N} A_t(i,j)_{l_j=B}$$



Triangle affinities

$\forall i, j = 1, \ldots, N,$

$$dist_G(i,j) \xrightarrow{GKF} A'_t(i,j) \xrightarrow[]{\substack{\text{linearly}\\ \text{rescale}}} A_t(i,j),$$

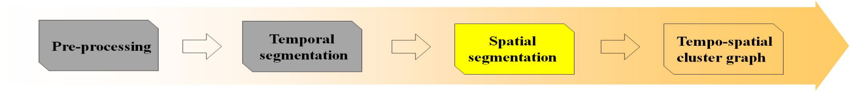
where $A'_t(i,j) \in [0\ 1]$ and $A_t(i,j) \in [-1\ 1]$.

Algorithm 2: *Adaptive-label ICM*

Init: l_i=1, i=1,..., n, L=1, A_t

- **repeat**
  - **for** i=1 to n **do**
    - **for** l=1 to L **do**
      - $A_l = \sum_{j=1}^{n} A_t(i,j)_{l_j=l}$
    - **end for**
    - **if** $\forall A_l < 0$ **then**
      - L=L+1
      - l_i=L
    - **else**
      - [$A_{l\text{-}max}$,l_{max}]=**max**(A_l)
      - l_i=l_{max}
    - **end if**
  - **end for**
- **until** L is unchanged
- **return** l_i, i=1,..., n

Deformation-based merging



Initial segmentation

Deformation-based merging

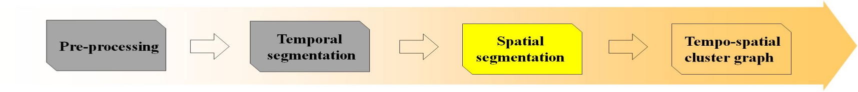
Average deformation of spatial cluster L_c, in the k-th temporal segment,

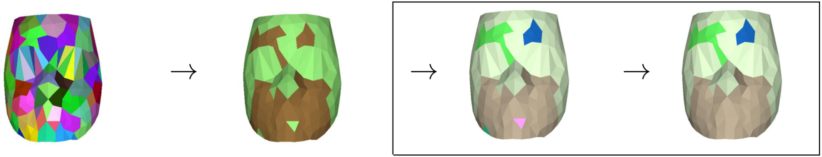
$$\overline{deform(c)^k} = \frac{\sum_{p\in[p_k,q_k],l_i=L_c} |s_i^p|}{\alpha_c}.$$

Threshold,

$$\rho_s(k) = \omega \cdot \overline{deform^k} = \omega \cdot \frac{\sum_{p\in[p_k,q_k]} |\mathbf{s}^\mathbf{p}|}{\sum_{c=1}^{L} \alpha_c}.$$

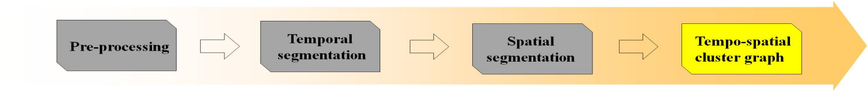
Post-processing





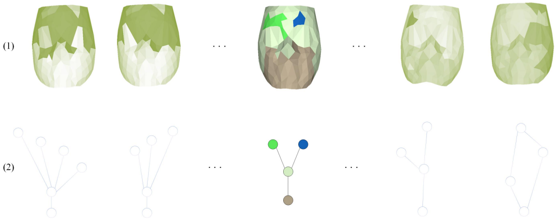
Spatial segmentation on the 10-th temporal segment of 'Facial-expression' animation.

Tempo-spatial cluster graph



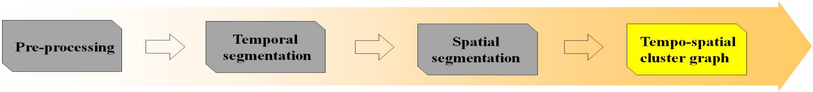
TSCG

- Graph representation
- Relaxed correspondence
- Transfer ratio



Tempo-spatial cluster graph of the 'Facial-expression' animation.

Tempo-spatial cluster graph

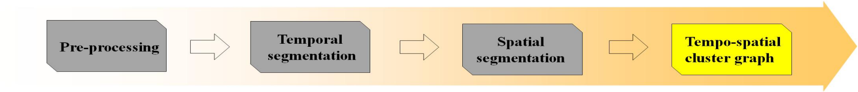


TSCG

- Graph representation
- Relaxed correspondence
- Transfer ratio

Tempo-spatial cluster graph of the ‘Facial-expression’ animation.

Tempo-spatial cluster graph

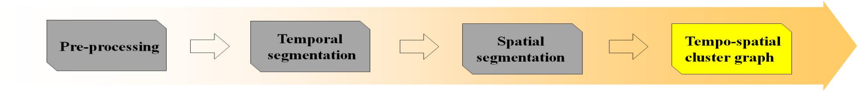


TSCG

- Graph representation
- Relaxed correspondence
- Transfer ratio

Tempo-spatial cluster graph of the ‘Facial-expression’ animation.

Tempo-spatial cluster graph

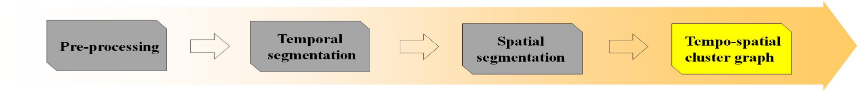


TSCG

- Graph representation
- Relaxed correspondence
- Transfer ratio

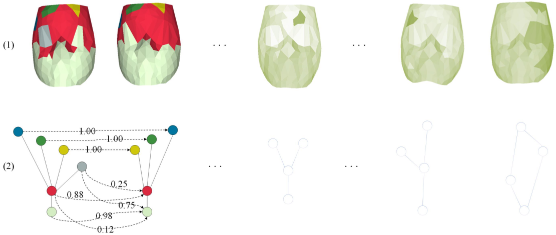
Tempo-spatial cluster graph of the 'Facial-expression' animation.

Tempo-spatial cluster graph



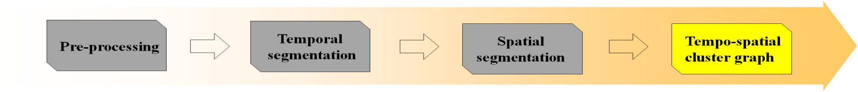
TSCG

- Graph representation
- Relaxed correspondence
- Transfer ratio



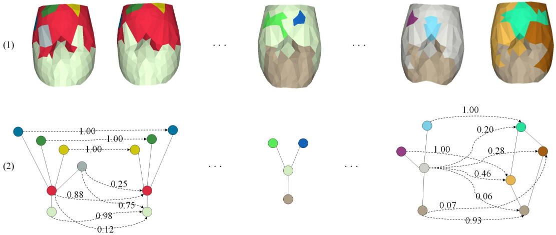
Tempo-spatial cluster graph of the ‘Facial-expression’ animation.

Tempo-spatial cluster graph



TSCG

- Graph representation
- Relaxed correspondence
- Transfer ratio



Tempo-spatial cluster graph of the ‘Facial-expression’ animation.

Results

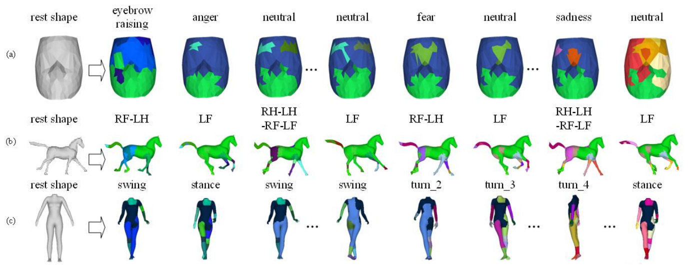
Timing

Time (sec) / Data	size # triangles	size # frames	Pre-processing	Temporal segmentation	Spatial segmentation
Facial expressions	286	1098	7.68	347.15	0.65
Galloping horse	8420	48	572.44	0.26	218.89
Walking woman	8590	279	656.82	11.14	228.64

Segmentation timings of animations.

Results

Tempo-spatial segmentation



(a) 'Facial-expression' animation, (b) 'Gallop-horse' animation (RF: right fore, LF: left fore, RH: right hint, LH: left hint.), (c) 'Walking-woman' animation.

The complete results

Conclusions

- We have introduced a method for temporal segmentation of mesh sequence.
- We develop geometric segmentation of mesh sequence.
- We devise an abstract representation of mesh sequence.
- Our segmentation results faithfully reflect the movements in the given mesh sequence.

Future works

- Improving of the relaxed correspondence algorithm.
- Comparison with other segmentation works.
- Consistent segmentation.
- Shape query.
- Compression of animations.

Acknowledgements

We would like to thank Frederic Larue and Olivier Genevaux for providing us with the facial motion capture data.
This work has been supported by the French national project SHARED (Shape Analysis and Registration of People Using Dynamic Data, No.10-CHEX-014-01).

köszönöm תודה! *děkuji*

mahalo 고맙습니다

thank you

merci 谢谢 *danke*

Ευχαριστώ شكرا

どうもありがとう *gracias*

References-I

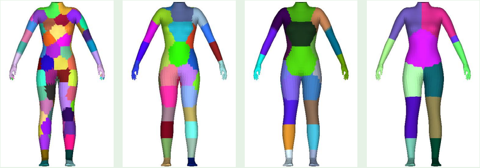
- Sattler M., Sarlette R., Klein R., Simple and Efficient Compression of Animation Sequences, Eurographics/ACM SIGGRAPH Symposium on Computer Animation, 2005.
- MAMOU K., ZAHARIA T., PRETEUX F.: A skinning approach for dynamic 3d mesh compression: Research articles. Comput. Animat. Virtual Worlds 17, 304(2006), 337-346.
- MAMOU K., ZAHARIA T., PRETEUX F.: The mpeg-4 standard for animated mesh compression. In ICIP 08, pages 2676-2679.
- MAMOU K., ZAHARIA T., PRETEUX F.: A DCT-based approach for dynamic 3D mesh compression. WSEAS Transactions on Information Science and Applications, Vol. 3(8), 2006, pages 1947-1954.
- Lee T.-Y., Wang Y.-S., Chen T.-G., Segmenting a Deforming Mesh into Near-Rigid Components, The Visual Computer, 22(9):729 - 739, 2006.
- Kalafatlar E. and Yemez Y., 3D Articulated Shape Segementation Using Motion Information, Int'l Conf. on Pattern Recognition, 3595–3598, 2010.
- CHEN X., GOLOVINSKIY A., FUNKHOUSER T.: A benchmark for 3D mesh segmentation. ACM Transactions on Graphics (SIGRRAPH) 28, 3 (2009).
- BENHABILES H., VANDEBORRE J.-P., LAVOUE G., DAOUDI M.: A framework for the objective evaluation of segmentation algorithms using a ground-truth of human segmented 3D models. In IEEE international Conference on Shape Modeling and Applications, 36-43 (2009).
- KALOGERAKIS E., HERTZMANN A., SINGH K.: Learning 3D Mesh Segmentation and Labeling. ACM Transactions on Graphics (SIGGRAPH) 29, 3 (2010).
- Seo H., Kim S., Cordier F., Choi J. and Hong K., Estimating Dynamic Skin Tension Lines in Vivo using 3D Scans, Computer-Aided Design (Elsevier), accepted for ACM Symposium on Solid and Physical Modeling 2012, Dijon, France.

References-II

- BENHABILES H., LAVOUE G., VANDERBORRE J.-P., and DAOUDI M.: Learning boundary edges for 3D-mesh segmentation, Computer Graphics Forum - Eurographics Association - Ed. Blackwell, volume 30, issue 8, 2170-2182, (2011).
- KOVAR L., GLEICHER M., PIGHIN F.: Motion graphs. ACM Transactions on Graphics (TOG) 21, 3 (2002), 473-482.
- KAHOL K., TRIPATHI P., PANCHANATHAN S.: Automated gesture segmentation from dance sequences. In 6th IEEE International Conference on Automatic Face and Gesture Recognition, Seoul Korea (2004).
- BARBI J., SAFONOVA A., PAN J. Y., FALOUTSOS C., HODGINS J. K., POLLARD N. S.: Segmenting motion capture data into distinct behaviors. In GI 04: Proceedings of the 2004 conference on Graphics Interface, 185-194 (2004).
- JANUS B., NAKAMURA Y.: Unsupervised Probabilistic Segmentation of Motion Data for Mimesis Modeling. In: IEEE International Conference on Advanced Robotics, 411-417, Seattle, USA (2005)
- WANG T., SHUN H., XU Y., and ZHENG N. 2001. Unsupervised Analysis of Human Gestures. IEEE Pacific Rim Conference on Multimedia 2001, 174-181.
- FOD A., MATARIC M. J., JENKINS O. C.: Automated derivation of primitives for movement classification. Auton. Robots 12, 1 (2002), 39-54.
- BAGON S.,GALUN M., Optimizing Large Scale Correlation Clustering, CoRR abs/111w.2903: 2011.
- O'Rourke J., Computational geometry column 35. SIGACT News, 30(2): 31-32, 1999.
- MITCHELL J. S. B.,MOUNT D. M., and PAPADIMITRIOU C. H.,The discrete geodesic problem. SIAM J. of Computing 16(4), 647668, 1987.

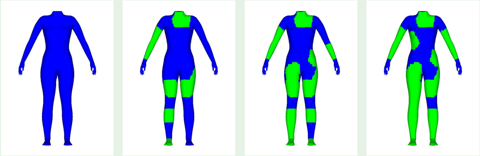
Discussions

Evaluation of δ_s



(o) δ_s=0.1 (p) δ_s=0.15 (q) δ_s=0.2 (r) δ_s=0.3

Evaluation of ω



(s) ω=0.3 (t) ω=0.5 (u) ω=0.7 (v) ω=0.9