#### Lecture 06:

# Learning-based Character Animation

Libin Liu

School of Intelligence Science and Technology Peking University







VCL @ PKU

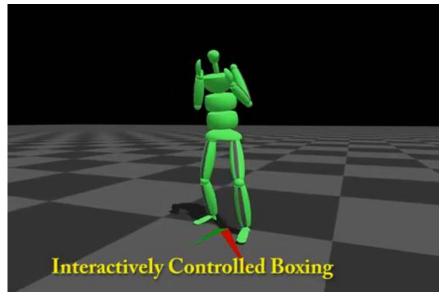
### Outline

- Recap: interactive character animation
  - Motion Graphs
  - Motion Matching
- Statistical Models of Human Motion
  - Principal Component Analysis
  - Gaussian Models
- Learning-based Models
  - .....

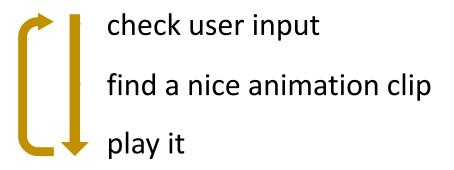
# Recap: Interactive Animation

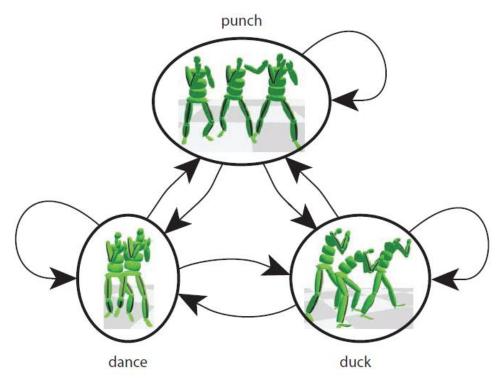
How to make a character respond to user command?

### How to create interactive animation?

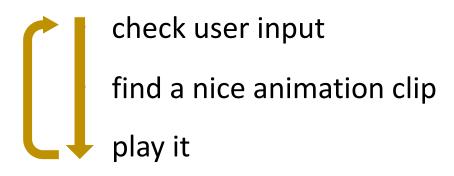


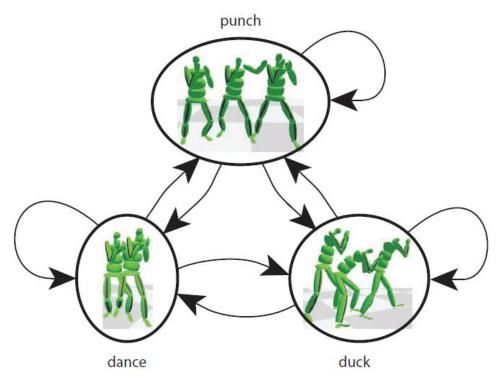
[Heck and Gleicher 2007, Parametric Motion Graphs]





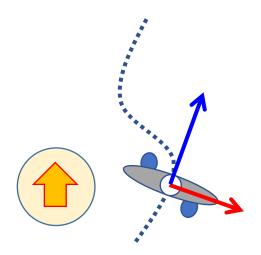
[Heck and Gleicher 2007, Parametric Motion Graphs]

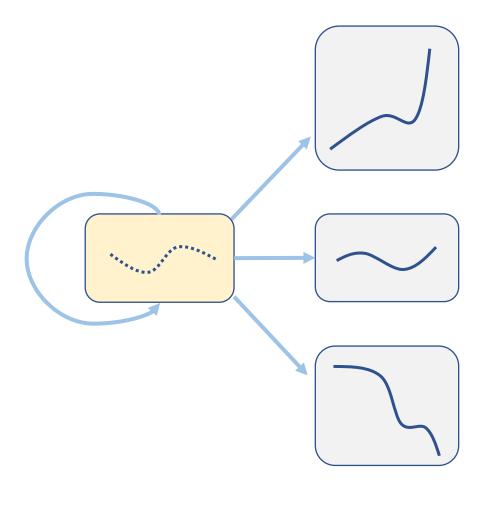


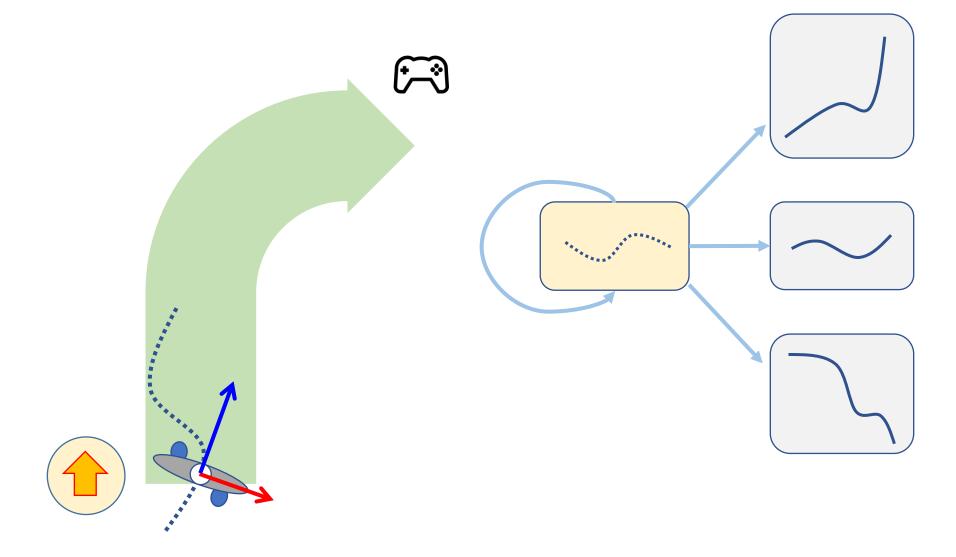


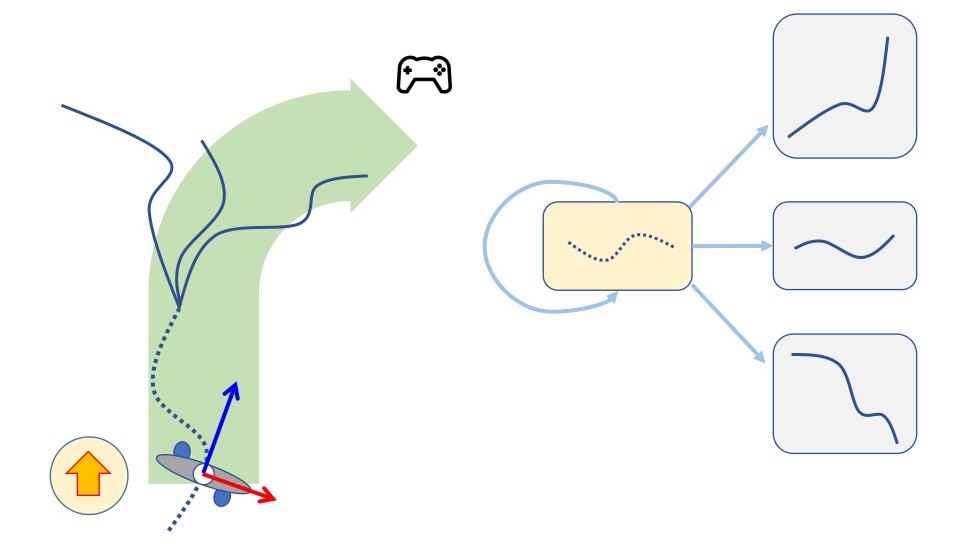
[Heck and Gleicher 2007, Parametric Motion Graphs]

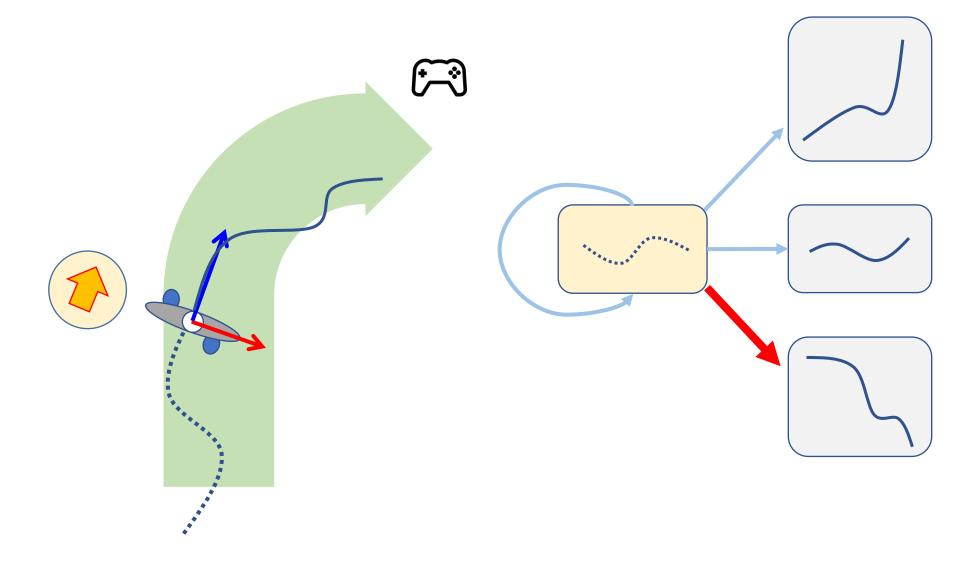
at the end of the current clip:
check user input
find a nice animation clip
play it





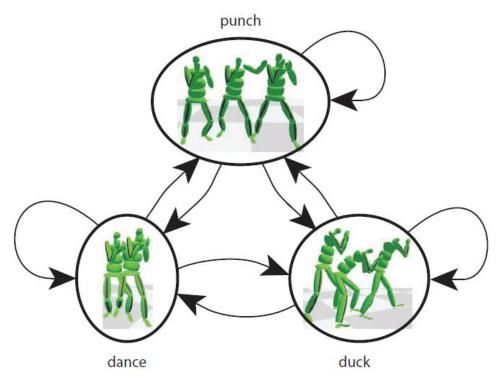




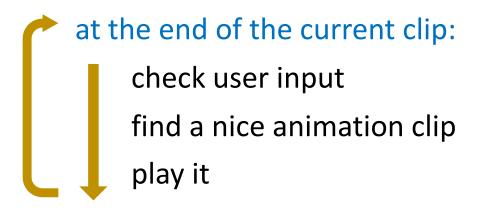




Motion Planning with Motion Graph and A\* https://www.youtube.com/watch?v=ekx0bXz25Pw



[Heck and Gleicher 2007, Parametric Motion Graphs]

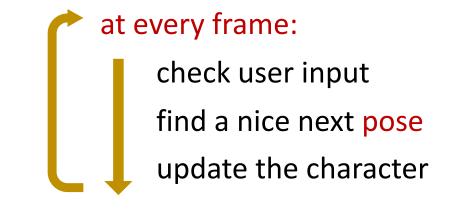


### Need a Faster Response?

Motion Graphs / State Machines

at the end of the current clip:
check user input
find a nice animation clip
play it





### Need a Faster Response?

Motion Graphs / State Machines

at the end of the current clip:

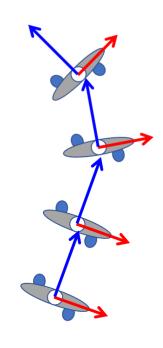
check user input find a nice animation clip play it

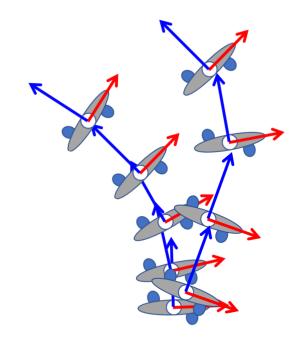
Motion Fields / Motion Matching

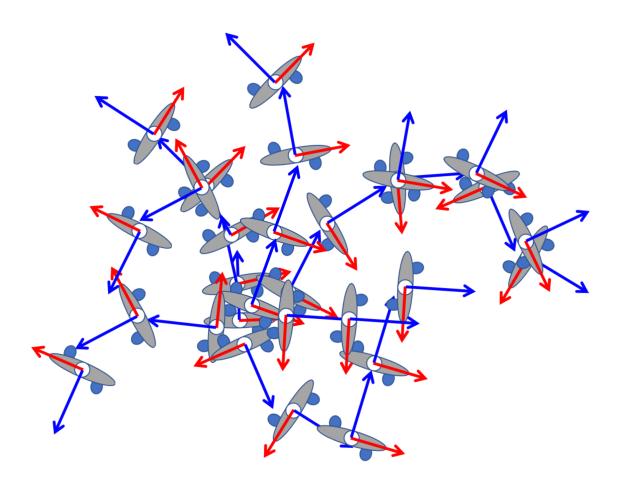
at every frame:

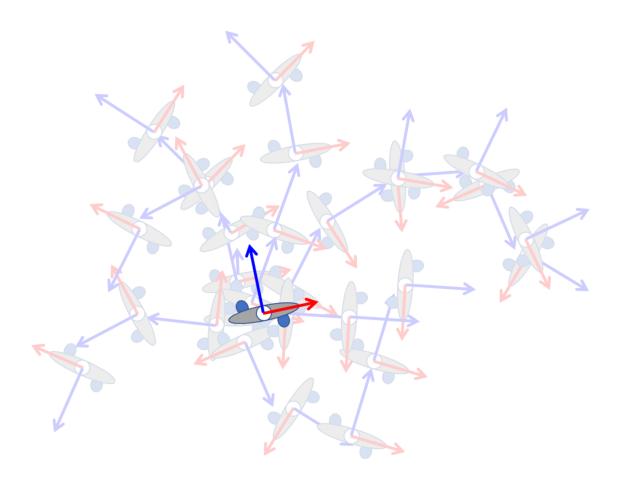
check user input find a nice next pose

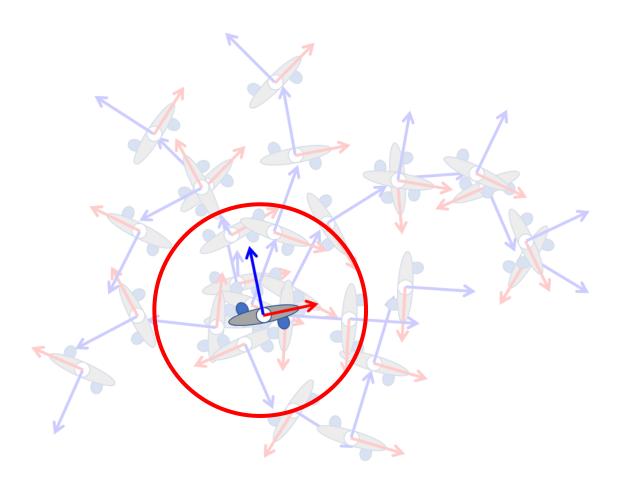
update the character

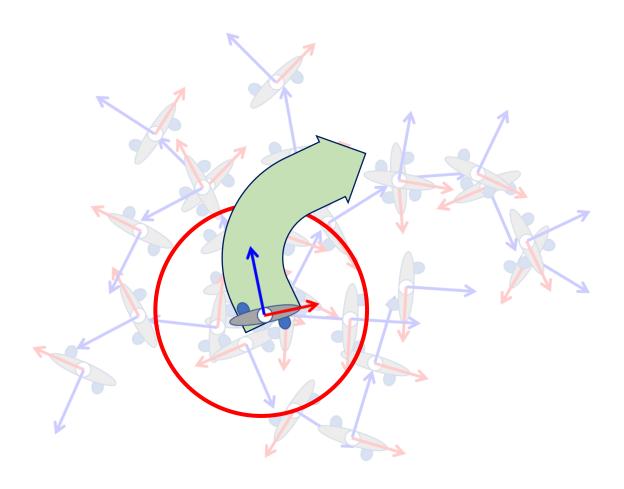










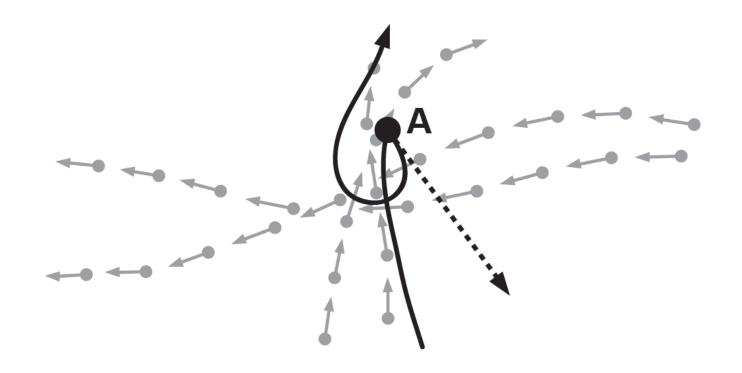


#### **Motion Fields for Interactive Character Locomotion**

Yongjoon Lee<sup>1,2\*</sup> Kevin Wampler<sup>1†</sup> Gilbert Bernstein<sup>1</sup> Jovan Popović<sup>1,3</sup> Zoran Popović<sup>1</sup>

<sup>1</sup>University of Washington <sup>2</sup>Bungie <sup>3</sup>Adobe Systems

\* SIGGRAPH 2010



#### **Motion Fields for Interactive Character Locomotion**

Yongjoon Lee<sup>1,2\*</sup>

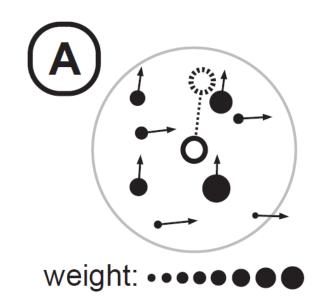
Kevin Wampler<sup>1†</sup>
<sup>1</sup>University of Washington

Gilbert Bernstein<sup>1</sup> on <sup>2</sup>Bungie

Jovan Popović<sup>1,3</sup>
<sup>3</sup>Adobe Systems

Zoran Popović<sup>1</sup>

\* SIGGRAPH 2010



#### **Motion Fields for Interactive Character Locomotion**

Yongjoon Lee<sup>1,2\*</sup>

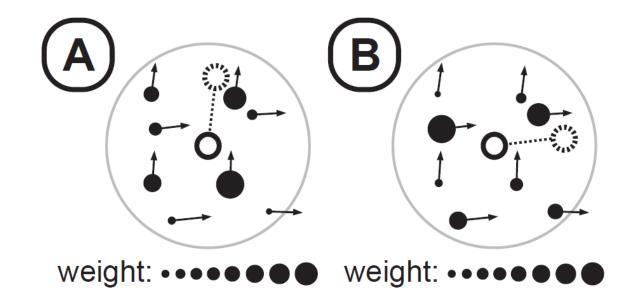
Kevin Wampler<sup>1†</sup>
<sup>1</sup>University of Washington

Gilbert Bernstein<sup>1</sup> on <sup>2</sup>Bungie

Jovan Popović<sup>1,3</sup>
<sup>3</sup>Adobe Systems

Zoran Popović<sup>1</sup>

\* SIGGRAPH 2010



#### Motion Fields for Interactive Character Locomotion

Yongjoon Lee<sup>1,2</sup>\*

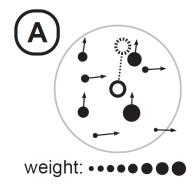
Kevin Wampler<sup>1†</sup>
<sup>1</sup>University of Washington

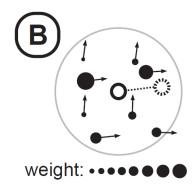
Gilbert Bernstein<sup>1</sup>
n <sup>2</sup>Bungie

Jovan Popović<sup>1,3</sup>
<sup>3</sup>Adobe Systems

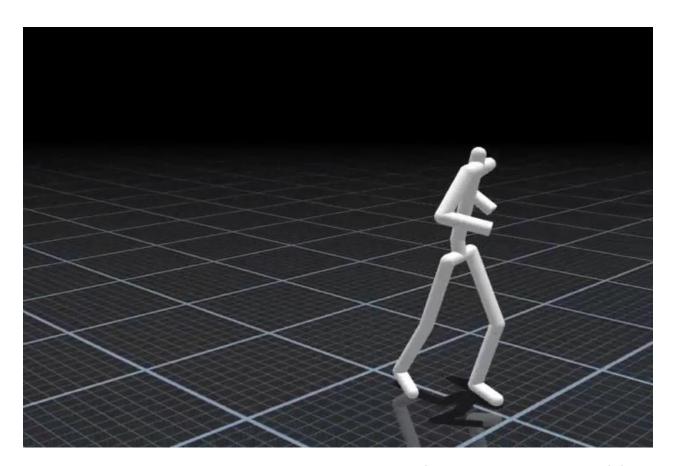
Zoran Popović<sup>1</sup>

\* SIGGRAPH 2010





- check user input
- find N nearest neighbors of the current state
- blend these neighbors according to user input
- update the character



Lee et al. 2010. Motion Fields

#### **Motion Fields**

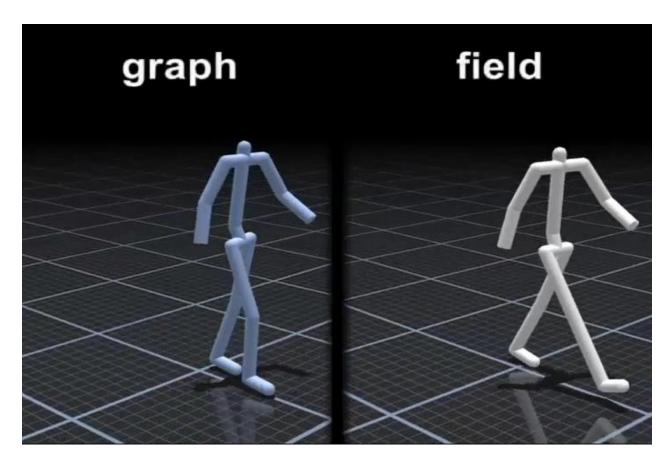
- check user input
- find N nearest neighbors of the current state
- blend these neighbors according to user input
- update the character



Lee et al. 2010. Motion Fields

**Motion Fields** 

- check user input
- find N nearest neighbors of the current state
- blend these neighbors according to user input
- update the character



Lee et al. 2010. Motion Fields

#### **Motion Fields**

- check user input
- find N nearest neighbors of the current state
- blend these neighbors according to user input
- update the character

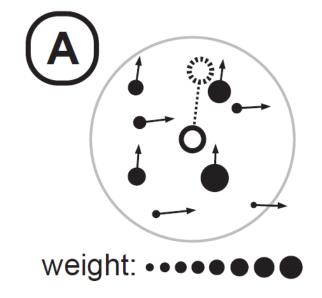
#### **Motion Fields**

- check user input
- find N nearest neighbors of the current state
- blend these neighbors according to user input
- update the character

**Motion Fields** 

#### at every frame:

- check user input
- find N nearest neighbors of the current state
- blend these neighbors according to user input
- update the character

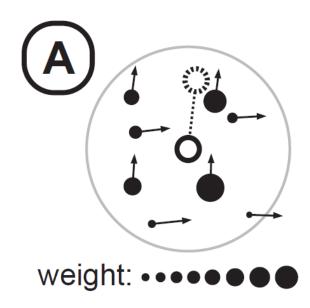


How?

**Motion Fields** 

#### at every frame:

- check user input
- find N nearest neighbors of the current state
- blend these neighbors according to user input
- update the character



How? Reinforcement learning...

**Motion Fields** 

#### at every frame:

- check user input
- find N nearest neighbors of the current state
- blend these neighbors according to user input
- update the character

**Motion Matching** 



- check user input
- find the nearest neighbors of the current state
   according to user input
- smoothly blend current pose to the nearest neighbor pose



We need a distance function / metric to define the nearest neighbor

We need a distance function / metric to define the nearest neighbor

$$next\_pose = \min_{i \in Dataset} ||x_{curr} - x_i||$$

x: feature vector

We need a distance function / metric to define the nearest neighbor

$$next\_pose = \min_{i \in Dataset} ||x_{curr} - x_i||$$

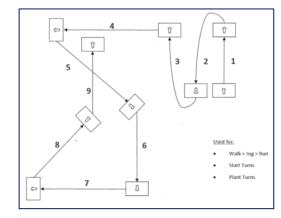
#### *x*: feature vector

A possible set of feature vectors:

- root linear/angular velocity
- position of end effectors w.r.t. root joint
- linear/angular velocity of end effectors w.r.t. root joint
- future heading position/orientation (e.g. in 0.5s, 1.0s, 1.5s, etc.)
- foot contacts
- .....

- We need a smooth motion
  - Only do the search every few frames
  - Smoothly blend current pose to the target pose
    - Inertialized blending (ref. <a href="https://www.theorangeduck.com/page/spring-roll-call">https://www.theorangeduck.com/page/spring-roll-call</a> by Daniel Holden)

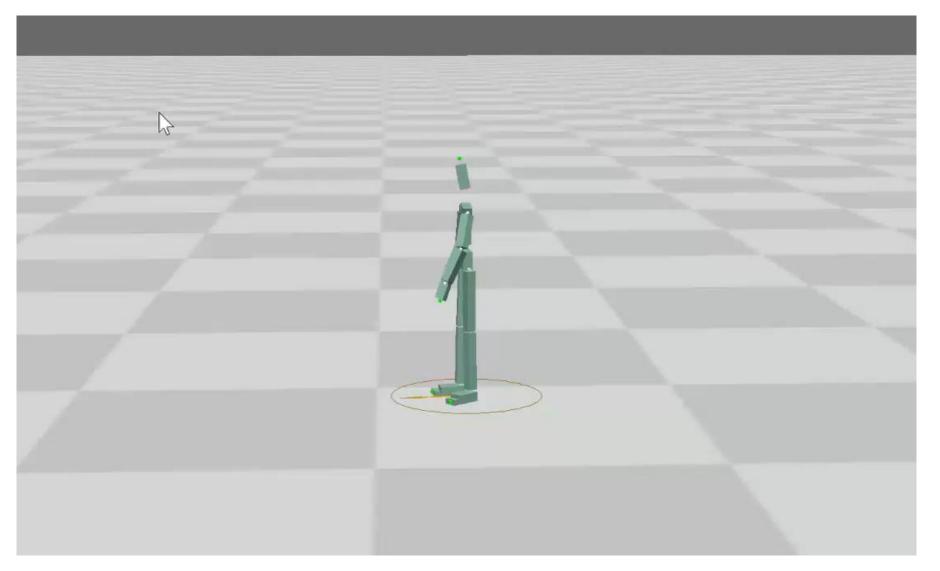
- We need a smooth motion
  - Only do the search every few frames
  - Smoothly blend current pose to the target pose
    - Inertialized blending (ref. <a href="https://www.theorangeduck.com/page/spring-roll-call">https://www.theorangeduck.com/page/spring-roll-call</a> by Daniel Holden)
- We need a good performance
  - An efficient data structure for searching
    - e.g. KD-tree
  - A efficient dataset
    - "Dance card"



## Motion Matching

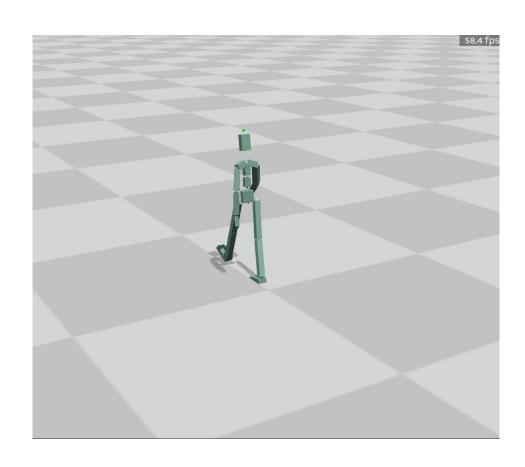


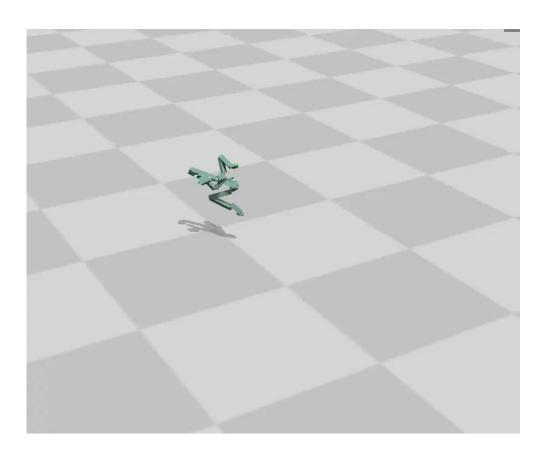
## Motion Matching



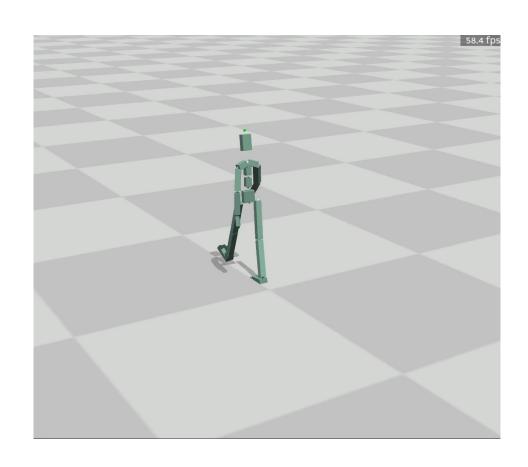
# Statistical Models of Human Motion

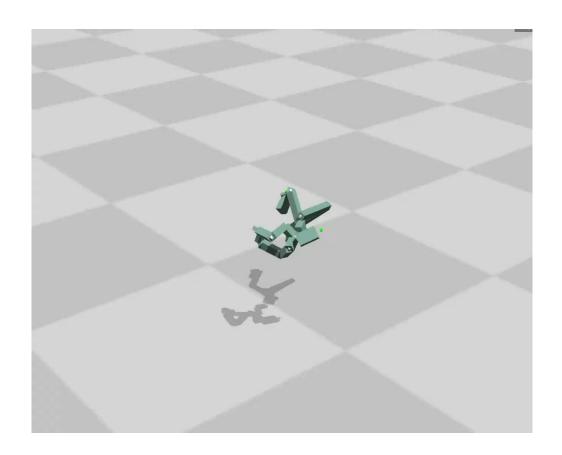
## What is a natural-looking motion?



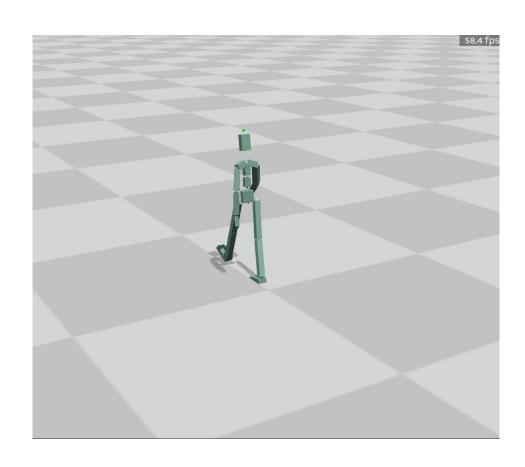


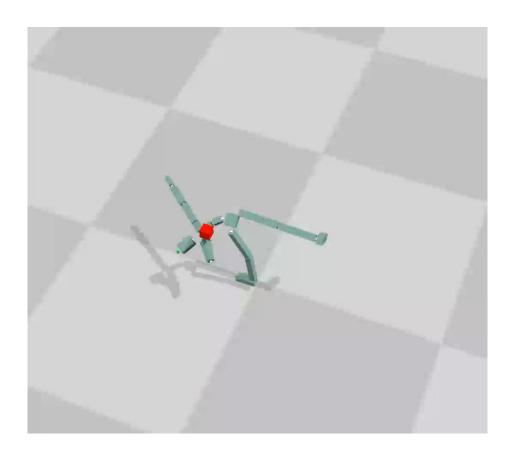
## What is a natural-looking motion?



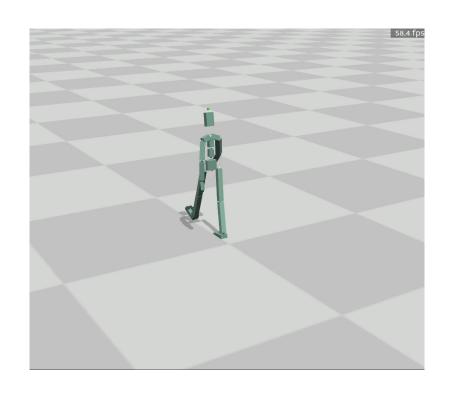


## What is a natural-looking motion?



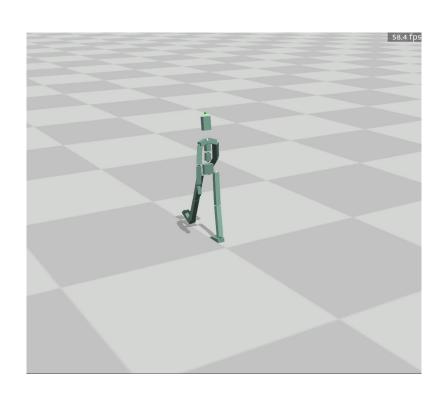


## The Low-dimensionality of Human Motions

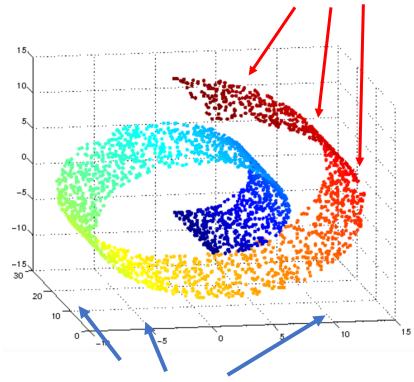


- Coordinated arm/leg movement
- Musculoskeletal structure
- Laws of physics
- •

## The Low-dimensionality of Human Motions

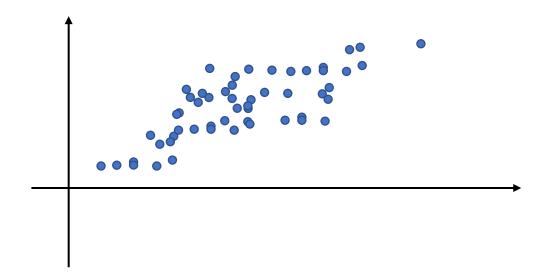


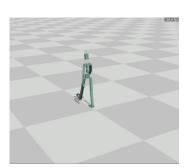
#### Where a natural motion locates



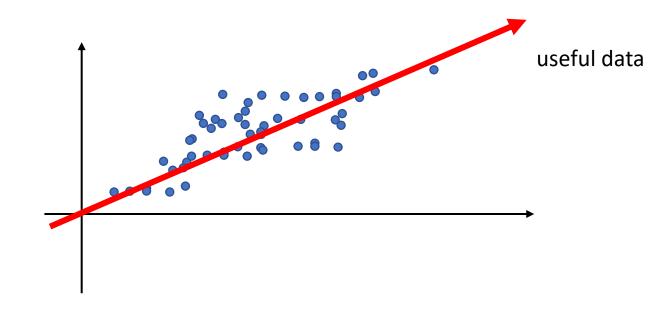
The entire "pose" space

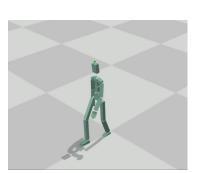
- A technique for
  - finding out the correlations among dimensions
  - dimensionality reduction



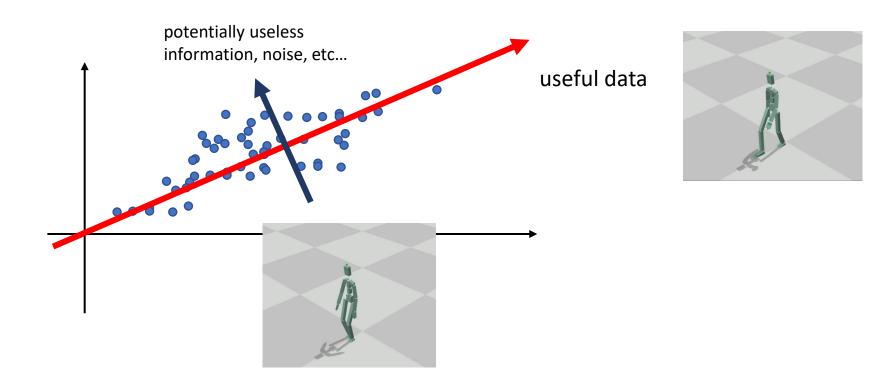


- A technique for
  - finding out the correlations among dimensions
  - dimensionality reduction

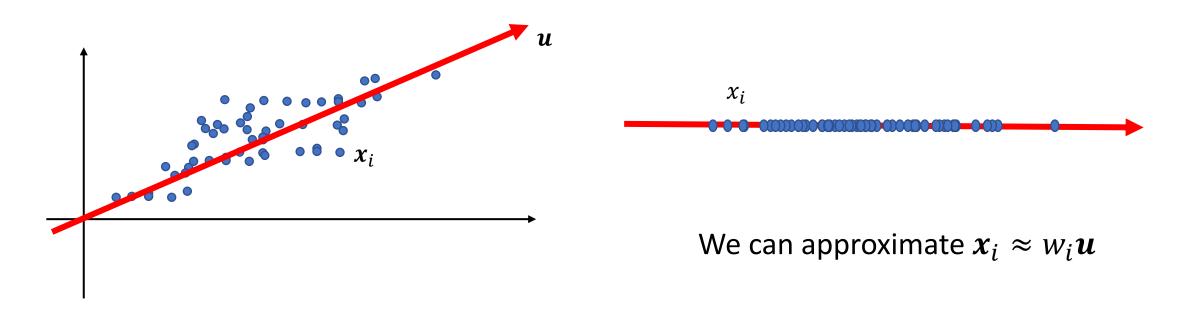


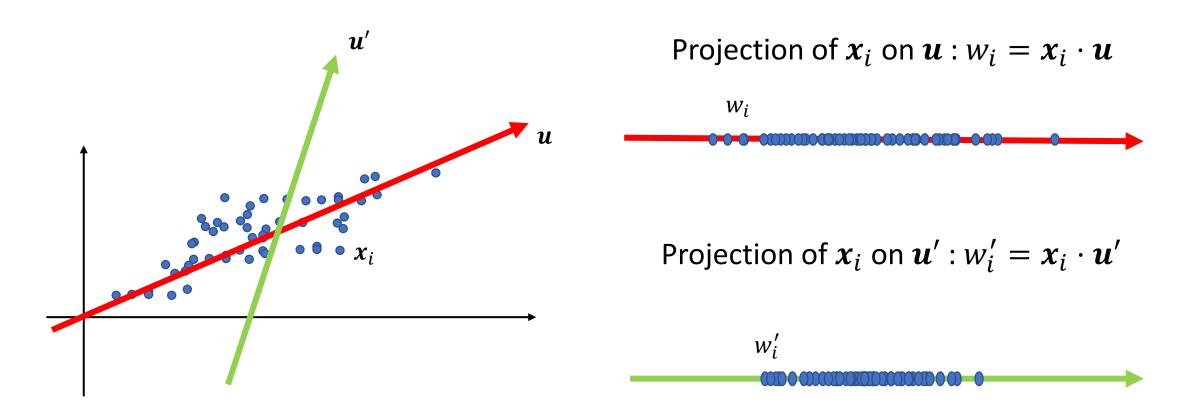


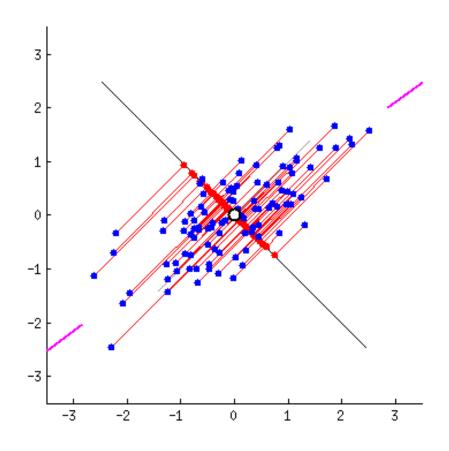
- A technique for
  - finding out the correlations among dimensions
  - dimensionality reduction

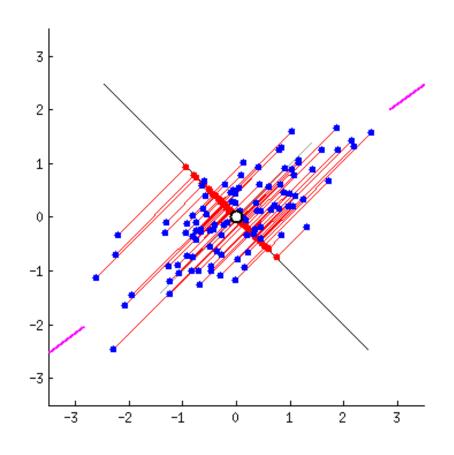


Projection of  $x_i$  on  $u: w_i = x_i \cdot u$ 









Find a direction  $\boldsymbol{u}$  such that  $\|\boldsymbol{u}\|=1$ , and the projections of  $\{\boldsymbol{x}_i\}$  on  $\boldsymbol{u}:w_i=\boldsymbol{x}_i\cdot\boldsymbol{u}$  have the maximal variance:

$$\frac{1}{N} \sum_{i} (w_i - \overline{w})^2$$

Find a direction u such that ||u|| = 1, and the projections of  $\{x_i\}$  on  $u: w_i = x_i \cdot u$ have the maximal variance:

$$\det X = \begin{bmatrix} (\boldsymbol{x}_0 - \overline{\boldsymbol{x}})^T \\ (\boldsymbol{x}_1 - \overline{\boldsymbol{x}})^T \\ \dots \\ (\boldsymbol{x}_N - \overline{\boldsymbol{x}})^T \end{bmatrix}$$

$$\frac{1}{N}\sum_{i}(w_{i}-\overline{w})^{2}$$

Find a direction  $\boldsymbol{u}$  such that  $\|\boldsymbol{u}\|=1$ , and the projections of  $\{x_i\}$  on  $u: w_i = x_i \cdot u$ have the maximal variance:

$$\det X = \begin{bmatrix} (\boldsymbol{x}_0 - \overline{\boldsymbol{x}})^T \\ (\boldsymbol{x}_1 - \overline{\boldsymbol{x}})^T \\ \dots \\ (\boldsymbol{x}_N - \overline{\boldsymbol{x}})^T \end{bmatrix}$$



$$\frac{1}{N}\sum_{i}(w_{i}-\overline{w})^{2}$$

It can be proved that  $oldsymbol{u}$  is an eigenvector of  $X^TX$  corresponds to the largest eigenvalue





Find a direction  $\boldsymbol{u}$  such that  $\|\boldsymbol{u}\|=1$ , and the projections of  $\{x_i\}$  on  $u: w_i = x_i \cdot u$ have the maximal variance:

$$\det X = \begin{bmatrix} (\boldsymbol{x}_0 - \overline{\boldsymbol{x}})^T \\ (\boldsymbol{x}_1 - \overline{\boldsymbol{x}})^T \\ \dots \\ (\boldsymbol{x}_N - \overline{\boldsymbol{x}})^T \end{bmatrix}$$



$$\frac{1}{N}\sum_{i}(w_{i}-\overline{w})^{2}$$

It can be proved that  $oldsymbol{u}$  is an eigenvector of  $X^TX$  corresponds to the largest eigenvalue



Note: we can approximate  $x_i \approx \overline{x} + w_i u$ 

• Given a dataset  $\{x_i\}$ ,  $x_i \in \mathbb{R}^N$ , then PCA gives

$$x_i = \overline{x} + \sum_{k=1}^n w_{i,k} u_k$$

- $oldsymbol{u_k}$  is the k-th principal component
  - A direction in  $\mathbb{R}^N$  along which the projection of  $\{x_i\}$  has the k-th maximal variance
- $w_{i,k} = (x_i \overline{x}) \cdot u_k$  is the score of  $x_i$  on  $u_k$

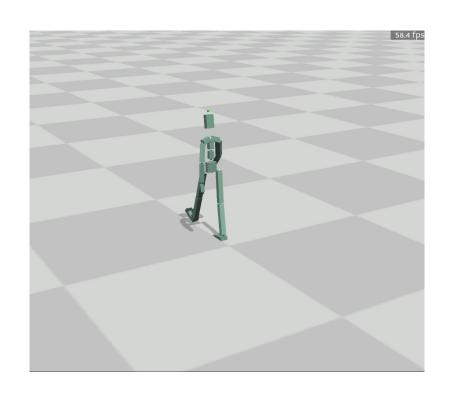
• Given a dataset  $\{x_i\}, x_i \in \mathbb{R}^N$ , the PCA can be computed by apply eigen decomposition on the covariance matrix

$$\Sigma = X^T X = U \begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_N^2 \end{bmatrix} U^T$$

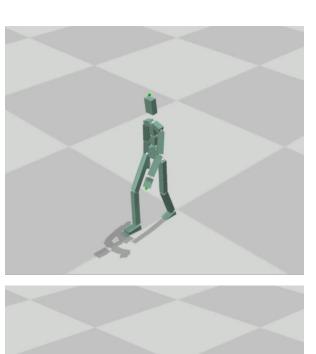
- $X = [x_0 \overline{x}, x_1 \overline{x}, ..., x_N \overline{x}]^T$
- $\sigma_i \ge \sigma_i \ge 0$  when i < j, corresponds to the Explained Variance
- $U = [u_1, u_2, ..., u_N]$

X

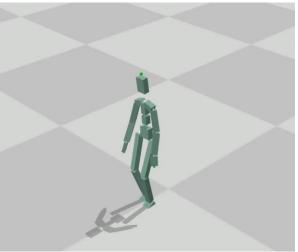
## PCA of Walking



 $x_i$ : joint rotations









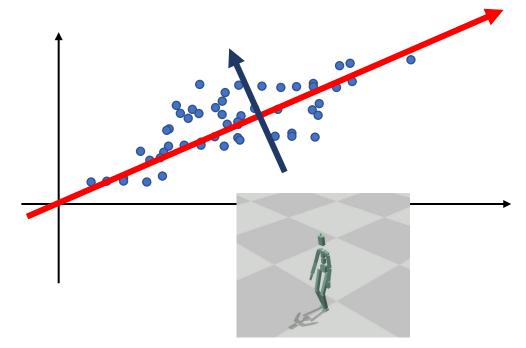
• Given a dataset  $\{x_i\}$ ,  $x_i \in \mathbb{R}^N$ , then PCA gives

$$x_i = \overline{x} + \sum_{k=1}^n w_{i,k} u_k$$

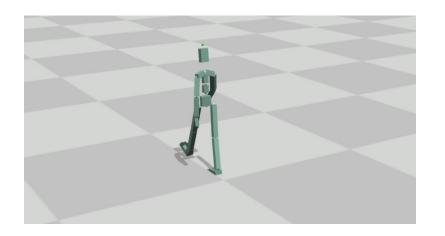
- $oldsymbol{u_k}$  is the k-th principal component
  - A direction in  $\mathbb{R}^N$  along which the projection of  $\{x_i\}$  has the k-th maximal variance
- $w_{i,k} = (x_i \overline{x}) \cdot u_k$  is the score of  $x_i$  on  $u_k$

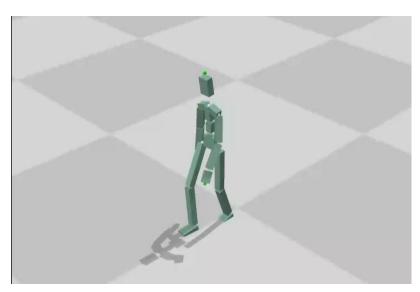
$$x_i = \overline{x} + \sum_{k=1}^n w_{i,k} u_k$$

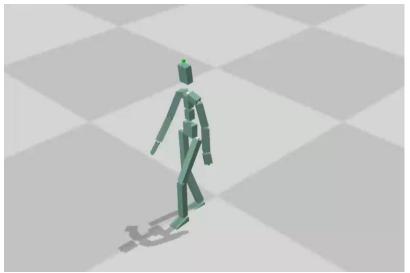




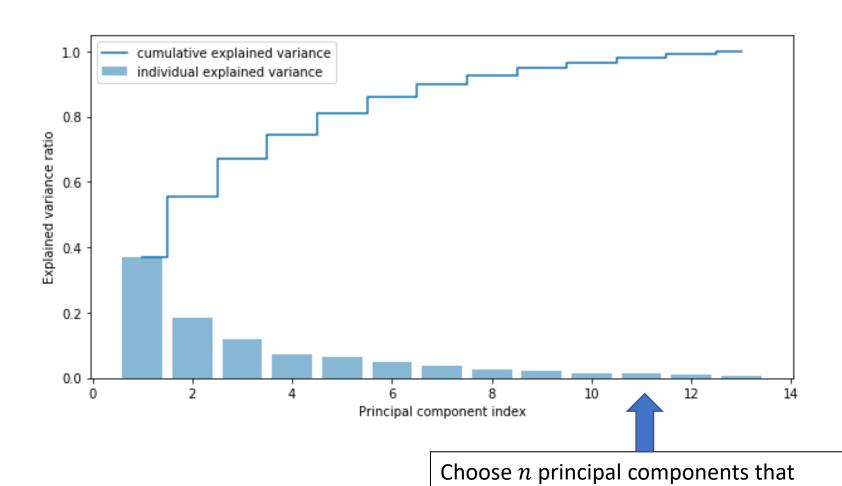
## PCA of Walking



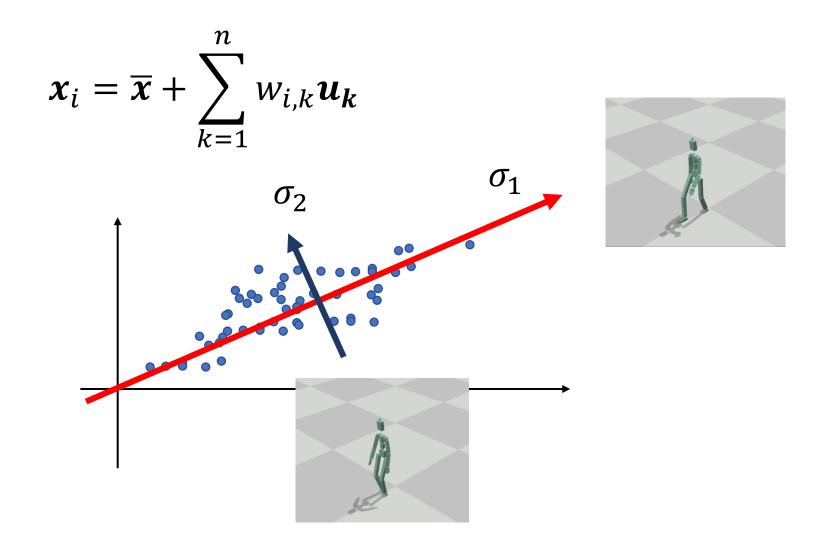




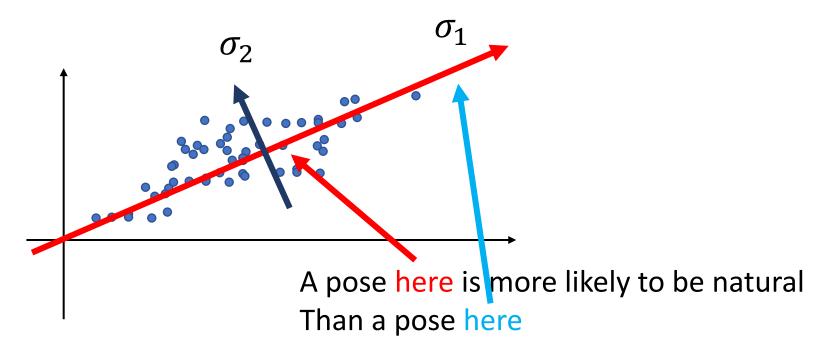




explains enough (e.g. 95%) of the variance

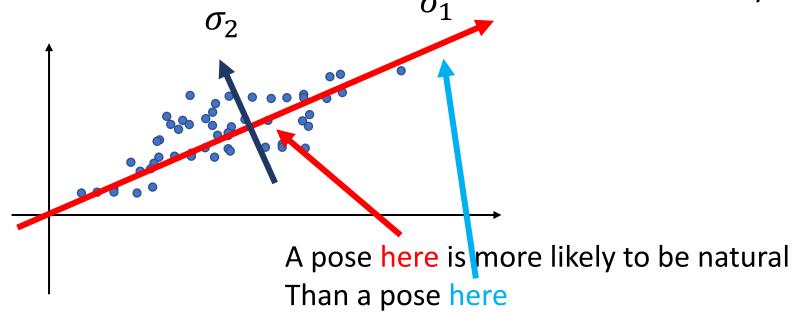


$$x_i = \overline{x} + \sum_{k=1}^n w_{i,k} u_k$$



$$x_i = \overline{x} + \sum_{k=1}^n w_{i,k} u_k$$

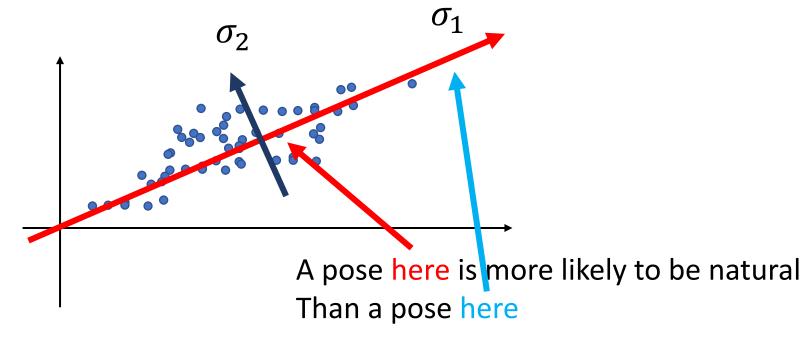
a pose  $x_i$  with smaller  $\sum_{k} \left(\frac{w_{i,k}}{\sigma_k}\right)^2$  is more likely to be a good pose



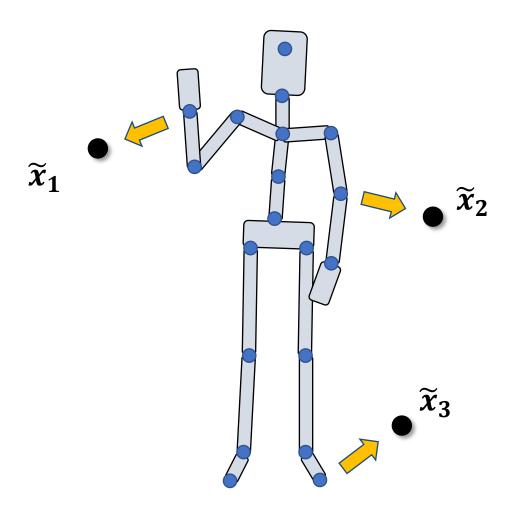
$$x = \overline{x} + \sum_{k=1}^{n} w_k u_k$$

a pose x with smaller  $\sum_{k} \frac{\left((x-\overline{x})\cdot u_{k}\right)^{2}}{\sigma_{k}^{2}}$ 

is more likely to be a good pose



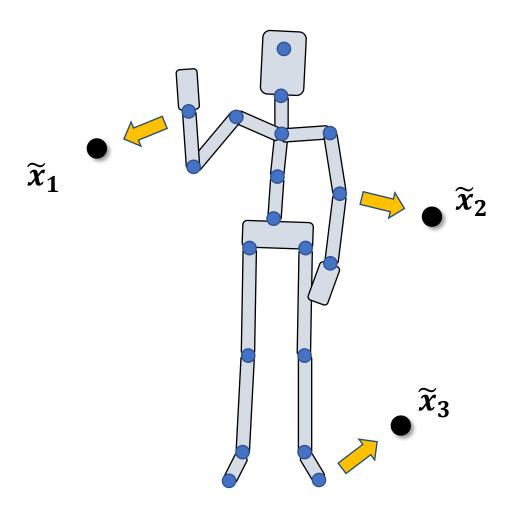
#### Character IK



$$F(\theta) = \frac{1}{2} \sum_{i} ||f_i(\boldsymbol{\theta}) - \widetilde{\boldsymbol{x}}_i||_2^2 + \frac{\lambda}{2} ||\boldsymbol{\theta}||_2^2$$

$$\boldsymbol{\theta} = (\boldsymbol{t}_0, R_0, R_1, R_2, \dots)$$

#### Character IK with a Reference Pose

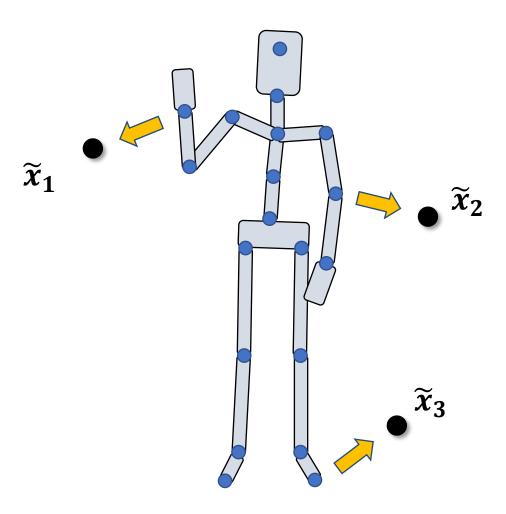


$$F(\theta) = \frac{1}{2} \sum_{i} \|f_i(\boldsymbol{\theta}) - \widetilde{\boldsymbol{x}}_i\|_2^2$$

$$+\frac{\lambda}{2}\|\boldsymbol{\theta}-\boldsymbol{\theta_0}\|_2^2$$

$$\boldsymbol{\theta} = (\boldsymbol{t}_0, R_0, R_1, R_2, \dots)$$

#### Character IK with a Motion Prior

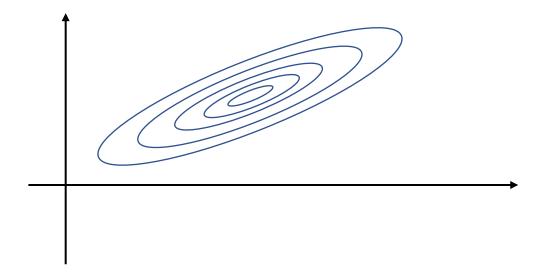


$$F(\theta) = \frac{1}{2} \sum_{i} ||f_i(\boldsymbol{\theta}) - \widetilde{\boldsymbol{x}}_i||_2^2$$

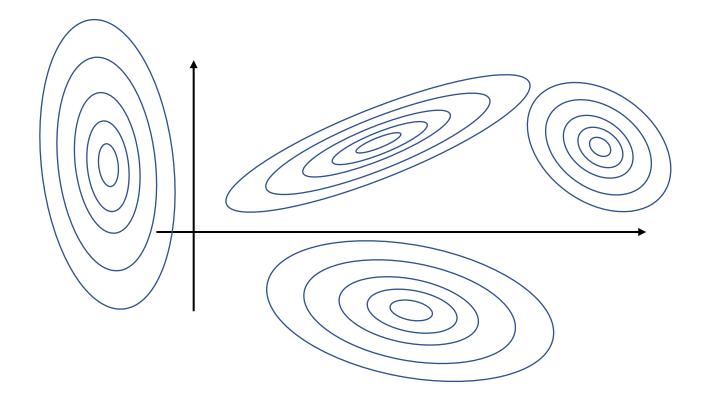
$$+\frac{w}{2}\sum_{k}\left(\frac{(\boldsymbol{\theta}-\overline{\boldsymbol{\theta}})\cdot\boldsymbol{u}_{k}}{\sigma_{k}}\right)^{2}$$

$$\boldsymbol{\theta} = (t_0, R_0, R_1, R_2, \dots)$$

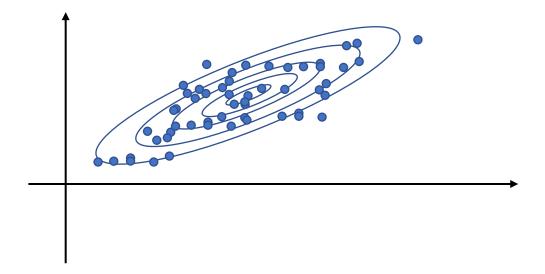
p(x): probability that x is a natural pose



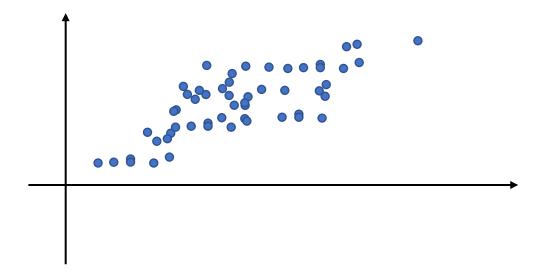
p(x): probability that x is a natural pose



p(x): probability that x is a natural pose a set of data points  $\{x_i\} \sim p(x)$ 



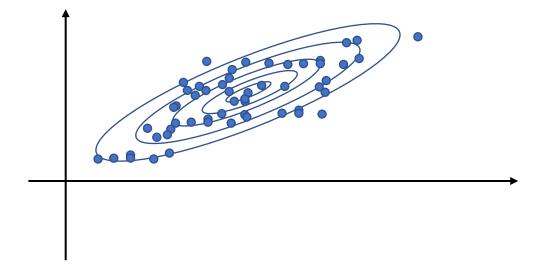
Given a dataset of mocap poses  $\{x_i\}$ 



### Data Distribution

Given a dataset of mocap poses  $\{x_i\}$ 

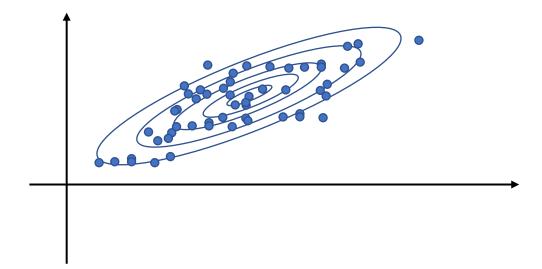
How to find p(x)?



### Gaussian Distribution

Dataset  $\{x_i\}$ 

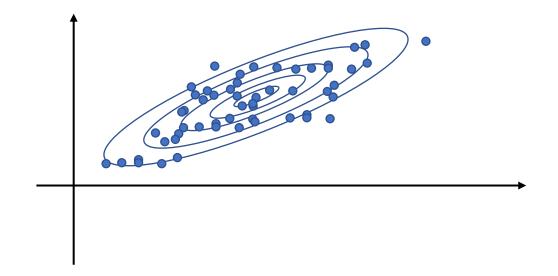
$$p(\mathbf{x}) = \mathcal{N}(\mu_i, \Sigma_i) = \frac{1}{\sqrt{(2\pi)^k |\Sigma|}} e^{-\frac{1}{2}(\mathbf{x} - \overline{\mathbf{x}})^T \Sigma^{-1} (\mathbf{x} - \overline{\mathbf{x}})}$$



### Gaussian Distribution

Dataset  $\{x_i\}$ 

$$p(\mathbf{x}) = \mathcal{N}(\mu_i, \Sigma_i) = \frac{1}{\sqrt{(2\pi)^k |\Sigma|}} e^{-\frac{1}{2}(\mathbf{x} - \overline{\mathbf{x}})^T \Sigma^{-1} (\mathbf{x} - \overline{\mathbf{x}})}$$



Maximum Likelihood Estimators (MLE):

$$\overline{x} = \frac{1}{N} \sum_{i} x_{i}$$

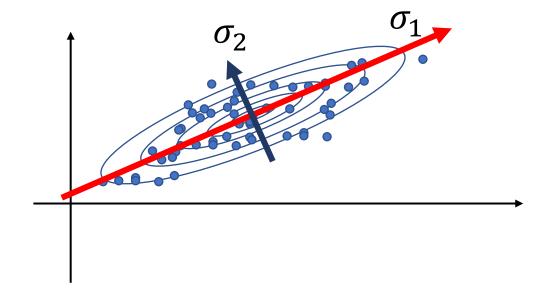
$$\Sigma = \frac{1}{N} X^T X$$



### PCA and Gaussian Distribution

Dataset  $\{x_i\}$ 

$$p(\mathbf{x}) = \mathcal{N}(\mu_i, \Sigma_i) = \frac{1}{\sqrt{(2\pi)^k |\Sigma|}} e^{-\frac{1}{2}(\mathbf{x} - \overline{\mathbf{x}})^T \Sigma^{-1} (\mathbf{x} - \overline{\mathbf{x}})}$$



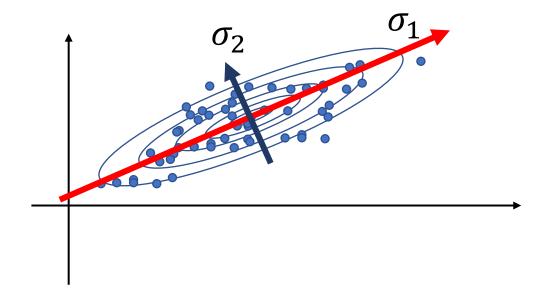
$$\Sigma = X^T X = U \begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_N^2 \end{bmatrix} U^T$$

$$x - \overline{x} = \sum_{k=1}^{n} w_k u_k$$

### PCA and Gaussian Distribution

Dataset  $\{x_i\}$ 

$$p(\mathbf{x}) = \mathcal{N}(\mu_i, \Sigma_i) = \frac{1}{\sqrt{(2\pi)^k |\Sigma|}} e^{-\frac{1}{2}(\mathbf{x} - \overline{\mathbf{x}})^T \Sigma^{-1} (\mathbf{x} - \overline{\mathbf{x}})}$$

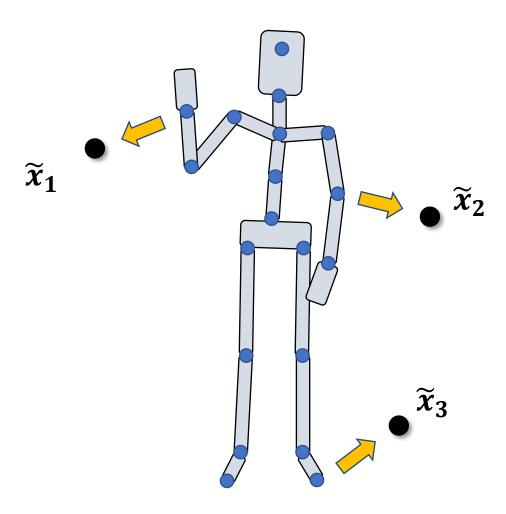




$$p(\mathbf{x}) = \prod_{k} \frac{1}{\sigma_k \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{w_k}{\sigma_k}\right)^2}$$

$$w_k = (x - \overline{x}) \cdot u_k$$

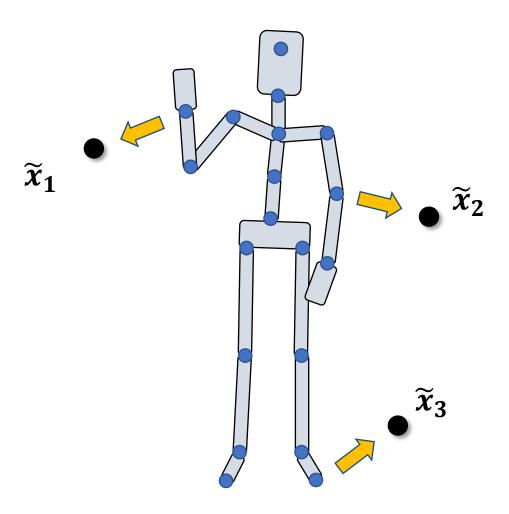
### Character IK with a Motion Prior



$$F(\theta) = \frac{1}{2} \sum_{i} ||f_{i}(\theta) - \widetilde{x}_{i}||_{2}^{2}$$
$$+ \frac{w}{2} \sum_{i} \left( \frac{(\theta - \overline{\theta}) \cdot u_{k}}{\sigma_{k}} \right)^{2}$$

$$\boldsymbol{\theta} = (\boldsymbol{t}_0, R_0, R_1, R_2, \dots)$$

### Character IK with a Motion Prior

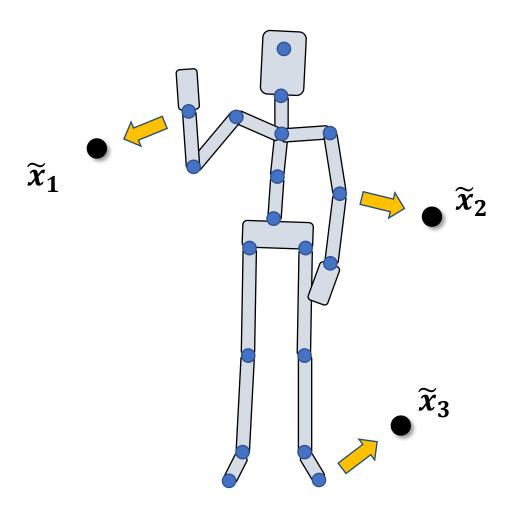


$$F(\theta) = \frac{1}{2} \sum_{i} \|f_i(\boldsymbol{\theta}) - \widetilde{\boldsymbol{x}}_i\|_2^2$$

$$-w\log\prod_{k}e^{-\frac{1}{2}\left(\frac{(\boldsymbol{\theta}-\overline{\boldsymbol{\theta}})\cdot\boldsymbol{u}_{k}}{\sigma_{k}}\right)^{2}}$$

$$\boldsymbol{\theta} = (t_0, R_0, R_1, R_2, \dots)$$

### Character IK with a Motion Prior



$$F(\theta) = \frac{1}{2} \sum_{i} ||f_i(\boldsymbol{\theta}) - \widetilde{\boldsymbol{x}_i}||_2^2$$

$$-w\log p(\theta)^{2} - \frac{1}{2} \left(\frac{(\theta - \overline{\theta}) \cdot u_{k}}{\sigma_{k}}\right)^{2}$$

$$\boldsymbol{\theta} = (t_0, R_0, R_1, R_2, \dots)$$

Given a motion prior p(x) learned from a set of data points  $D = \{x_i\}$ , Synthesize a motion x that minimize the objective

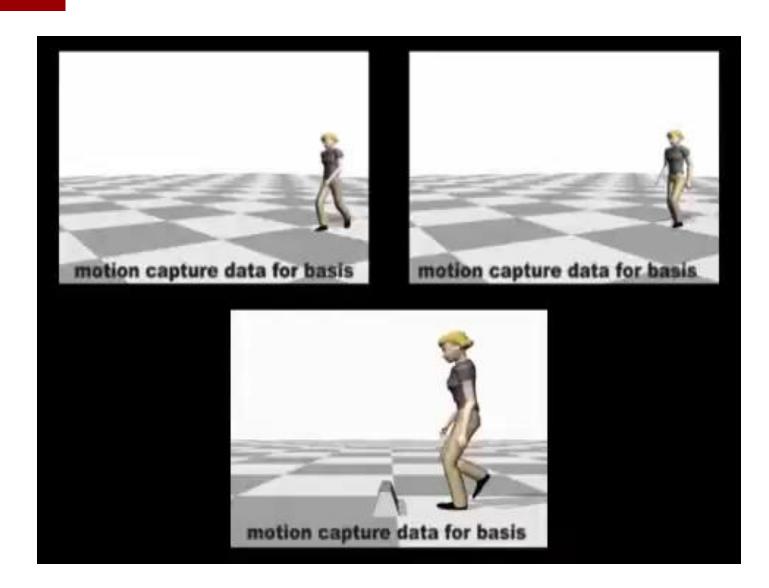
$$F(x) = f(x) - w \log p(x)$$

Note: x can represent a pose  $\theta$  or a motion clip  $\rightarrow$  a sequence of poses  $\{\theta_t\}$  or any features of a motion  $\rightarrow$  e.g.  $w_k$  in PCA

Given a motion prior p(x) learned from a set of data points  $D = \{x_i\}$ , Synthesize a motion x that minimize the objective

$$F(x) = f(x) - w \log p(x)$$

IK f(x) Keyframes
User control **Environment constraints** 



### Synthesizing Physically Realistic Human Motion in Low-Dimensional, Behavior-Specific Spaces

Alla Safonova

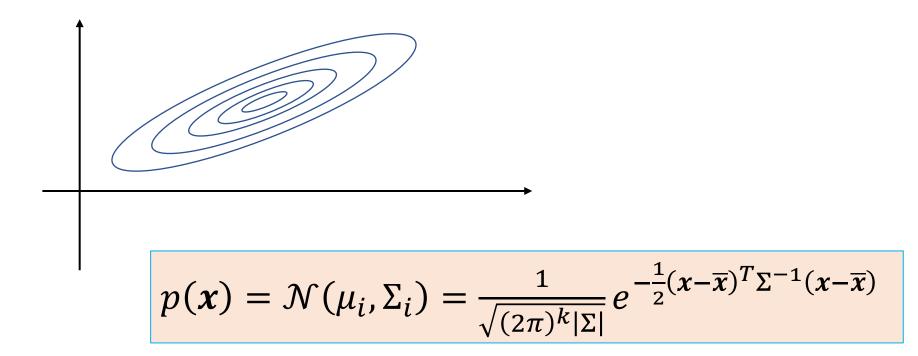
Jessica K. Hodgins

Nancy S. Pollard

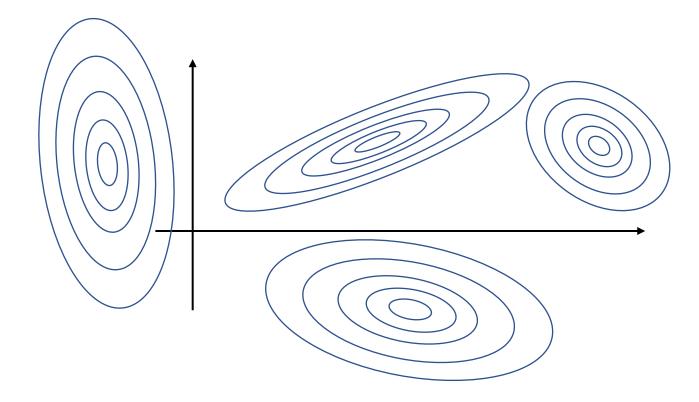
School of Computer Science Carnegie Mellon University \*

\*SIGGRAPH 2004

p(x): motion prior



p(x): motion prior

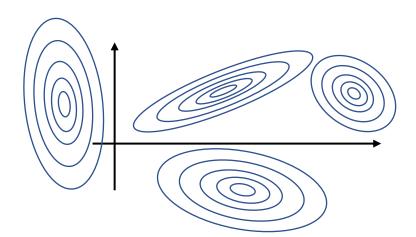


p(x): motion prior

#### Interactive Generation of Human Animation with Deformable Motion Models

Jianyuan Min Texas A&M University Yen-Lin Chen Texas A&M University Jinxiang Chai Texas A&M University

\* SIGGRAPH 2009



#### Gaussian Mixture Models (GMM)

$$p(\mathbf{x}) = \sum_{i} \phi_{i} \mathcal{N}(\mu_{i}, \Sigma_{i})$$

Interactive Generation of Human Animation with Deformable Motion Models

> Jianyuan Min Yen-Lin Chen Jinxiang Chai Texas A&M University

> > Min et al. 2009

### p(x): motion prior

#### Continuous Character Control with Low-Dimensional Embeddings

Zoran Popović<sup>2</sup> Sergey Levine<sup>1</sup> Jack M. Wang<sup>1</sup> Alexis Haraux<sup>1</sup> Vladlen Koltun<sup>1</sup> <sup>1</sup>Stanford University <sup>2</sup> University of Washington

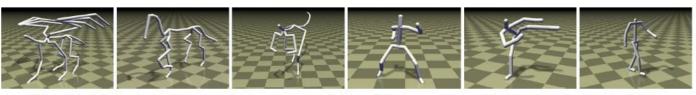
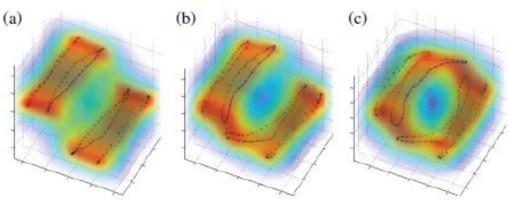


Figure 1: Character controllers created using our approach: animals, karate punching and kicking, and directional walking.

\* SIGGRAPH 2012

#### Gaussian Process Latent Variable Model (GPLVM)



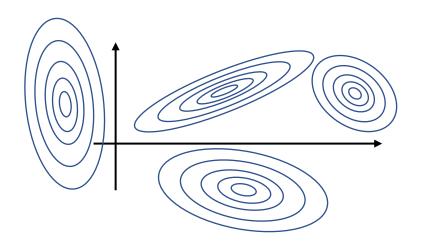
# Continuous Character Control with Low-Dimensional Embeddings

Sergey Levine<sup>1</sup> Jack M. Wang<sup>1</sup> Alexis Haraux<sup>1</sup> Zoran Popović<sup>2</sup> Vladlen Koltun<sup>1</sup>

<sup>1</sup>Stanford University <sup>2</sup>University of Washington

Levine et al. 2012

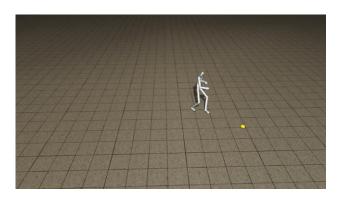
### p(x): motion prior



#### Neural networks...



[Starke et al 2020, Local Motion Phases for Learning Multi-Contact Character Movements]



[Lee et al 2019, Interactive Character Animation by Learning Multi-Objective Control]



[Henter et al. 2020, MoGlow: Probabilistic and Controllable Motion Synthesis Using Normalising Flows]



[Holden et al 2020, Learned Motion Matching]

# Questions?

