Kinematic Features for Action Recognition
Saad Ali & Mubarak Shah


General Idea

• Global Motion Dynamics for Action Recognition

Video
Optical Flow
Encoding Scheme
Kinematic Features
General Idea

PCA of Kinematic Features to produce dominant Kinematic Modes

Bag of Kinematic Model

Multiple Instance Learning

Kinematic Feature

• Divergence
• Vorticity
• Symmetric Flow
• Asymmetric Flow
• Invariants of Gradient Tensor
• Invariants of Rate of Rotation Tensor
• Invariants of Rate of Strain Tensor
Dominant Kinematic Trends

- Hypothesis: Dynamical information of optical flow field can be captured in terms of dominant kinematic trends.

- Dominant Kinematic Trends: Non-Linear Coherent Structure (NLCS, modes)

- Discovered by employing kinematic features as kinematic kernels in a PCA framework.

- Allows discovery of dominant modes in terms of dynamics, not energy.

Classification Methodology

- Multiple Instance based Learning Algorithm

- Video = Bag of NLCS

- Each bag is embedded into an instance based feature space

- Coordinates of the bag in the feature space are used for classification
Algorithmic Steps

Optical Flow Computation

- Block based cross correlation in frequency domain
- Peak in the correlation surface are used to compute the displacement
- Outliers are removed by adaptive local median filtering.
- Removed vectors are filled by interpolation from the neighboring velocity vectors
Optical Flow Computation

Kinematic Features

Divergence

• A scalar quantity, defined at a point \((x, t_i)\) in space and time as:

\[
\mathbf{f}^1(x, t_i) = \frac{\partial u(x, t_i)}{\partial x} + \frac{\partial v(x, t_i)}{\partial y}
\]

• Capture amount of local expansion in the flow field

• Useful for discriminating action involving motion of different body parts:
  – Hand Wave vs. Bend
Kinematic Features

Divergence

- Measure of local spin along the axis perpendicular to the plane of the flow field
- Computed at a point \( (x, t_i) \) as follows:

\[
f^2 (x, t_i) = \frac{\partial v(x, t_i)}{\partial x} - \frac{\partial u(x, t_i)}{\partial y}
\]
- Also measures rigidity in the flow
- Important for distinguishing between actions that involve articulated motion and the ones that do not

Vorticity (Curl)

- Measure of local spin along the axis perpendicular to the plane of the flow field
- Computed at a point \( (x, t_i) \) as follows:
Kinematic Features

Vorticity

Symmetric & Asymmetric Flow Fields

• Captures symmetry and asymmetry of human actions around the diagonal axis

• Raising Right Hand vs. Raising Both Hands

\[
\begin{align*}
    f^3(t_i) &= u(t_i) + u(t_i)^* \\
    f^5(t_i) &= u(t_i) - u(t_i)^* \\
    f^4(t_i) &= v(t_i) + v(t_i)^* \\
    f^6(t_i) &= v(t_i) - v(t_i)^*
\end{align*}
\]

Symmetric  \hspace{1cm}  Asymmetric
Kinematic Features
Symmetric Flow Fields

Kinematic Features
Asymmetric Flow Fields
Kinematic Features
Gradient Tensor

• Small scale structures are often present in a flow filed – Referred to as ‘eddies’

• These structures are characterized by large velocity gradients

• In a flow field representing human actions, eddies arise due to local motion of different limbs.

• Flow field gradients are a better measure of the local structure than the raw flow field itself.

\[
\begin{bmatrix}
\frac{\partial u(x, t_i)}{\partial x} & \frac{\partial u(x, t_i)}{\partial y} \\
\frac{\partial v(x, t_i)}{\partial x} & \frac{\partial v(x, t_i)}{\partial y}
\end{bmatrix}
\]

• Optical flow gradient tensor:

• Kinematic features derived from the gradient tensor are based on the concept of tensor invariants

• Tensor invariants: Scalar quantities which remain unchanged under full transformation group
Kinematic Features
Gradient Tensor

• Compute three principal invariants of the gradient tensor:

\[ P(x, t_i) = -\text{trace} \left( \nabla U(x, t_i) \right) \]

\[ Q(x, t_i) = \frac{1}{2} \left( P^2 - \text{trace} \left( \nabla U(x, t_i)^2 \right) \right) \]

\[ R(x, t_i) = -\det \left( \nabla U(x, t_i) \right) \]

\[ f^7(t_i) = Q(t_i) \quad f^8(t_i) = R(t_i) \]
Kinematic Features
Third Invariant - Gradient Tensor

Kinematic Features
Rate of Strain and Spin Tensors

- Obtain Rate of Strain and Rate of Spin tensors by decomposing Gradient tensor

\[ S(x, t_i) = \frac{1}{2} (\nabla U(x, t_i) + \nabla U(x, t_i)^*) \]

\[ O(x, t_i) = \frac{1}{2} (\nabla U(x, t_i) - \nabla U(x, t_i)^*) \]

- These two tensors are used as a measure of deformability which results due to the presence of the gradients in the flow field.
- Represents deviations from the rigid body motion
- We use the second and third principal invariants of \( S \) and only the third invariant of \( O \)

\[ f^8(t_i) = Q_s(t_i) \quad f^9(t_i) = R_s(t_i) \quad f^{10}(t_i) = Q_o(t_i) \]
Kinematic Features
Second Invariant – Rate of Strain Tensor

Kinematic Features
Third Invariant – Rate of Strain Tensor
Kinematic Features
Third Invariant – Rate of Spin Tensor

Kinematic Modes

- Perform PCA of the kinematic features of the optical flow.

- Provides a richer description of the hidden dynamics of the optical flow.
Comparison of Modes

The energy-containing regions of the optical flow in the top row do not reveal the finer characteristics of the optical flow visible in the kinematic modes in the bottom row.

Kinematic Modes

The dominant kinematic modes of kinematic features for the "bend" action.
Kinematic Modes

The dominant kinematic modes of kinematic features for the “bend” action.

---

Kinematic Modes

The dominant kinematic modes of kinematic features for the “bend” action.
The dominant kinematic modes of kinematic features for the "two hands wave" action

Multiple Instance-based Learning

• The action classifier is learned using multiple instance learning (MIL).

• In MIL each example is represented by several feature vectors called “instances”.

• In our case each action (example) is represented by several kinematic modes, where each kinematic mode is treated as an instance representing the action.

• That is each video is represented as a bag (collection) of kinematic modes
Multiple Instance-based Learning

(a) Action Videos

(b) Bag of Kinematic Modes

Use K-means classifier for classification
Multiple Instance-based Learning

Coordinate values of the 10 actions classes of the Weizmann action data set.

Experiments

• Weizman Action Dataset

• Ten actions performed by nine different actors:
  – Bend, Jumping Jack, Jump Forward, Jump in Place, Run, Side Gallop, Walk, Wave One Hand, Wave Two Hands

• 90 videos
Experiments

Comparison with Optical Flow

<table>
<thead>
<tr>
<th></th>
<th>1-Mode</th>
<th>2-Modes</th>
<th>3-Modes</th>
<th>4-Modes</th>
<th>5-Modes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optical Flow</td>
<td>74.2%</td>
<td>76.43%</td>
<td>82.77%</td>
<td>82.25%</td>
<td>85.8%</td>
</tr>
<tr>
<td>Kinematic Features</td>
<td>80.3%</td>
<td>89.68%</td>
<td>93.18%</td>
<td>95.75%</td>
<td>94.75%</td>
</tr>
</tbody>
</table>
KTH Dataset

Six Actions
Run
Jog
Walk
Boxing
Hand-Clapping
Hand Waveing

KTH Dataset

<table>
<thead>
<tr>
<th></th>
<th>Boxing</th>
<th>Hand-Clapping</th>
<th>Hand-Waving</th>
<th>Jogging</th>
<th>Running</th>
<th>Walking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boxing</td>
<td>88.5%</td>
<td>10.2%</td>
<td>1.3%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Hand-Clapping</td>
<td>5.35%</td>
<td>86.44%</td>
<td>8.21%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Hand-Waving</td>
<td>5.4%</td>
<td>7.43%</td>
<td>84.46%</td>
<td>1.12%</td>
<td>1.59%</td>
<td>0%</td>
</tr>
<tr>
<td>Jogging</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>86.21%</td>
<td>9.78%</td>
<td>4.003%</td>
</tr>
<tr>
<td>Running</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>6.1%</td>
<td>91.51%</td>
<td>2.39%</td>
</tr>
<tr>
<td>Walking</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>7.49%</td>
<td>3.4%</td>
<td>89.11%</td>
</tr>
</tbody>
</table>

81.5% 81.17% 62.96% 71.72% 95.33% 71.16% 87.7%
Feature Contribution

2 = Divergence, 3 = Vorticity, 4 = Symmetric Flow (both u and v components), 5 = Asymmetric Flow (both u and v components), 6 = Second Invariant of the Gradient Tensor, 7 = Third Invariant of the Gradient Tensor, 8 = Second Invariant of the Rate of Strain Tensor, 9 = Third Invariant of Rate of Strain Tensor, 10 = Third Invariant of Rate of Rotation Tensor,

Questions