The paper presents a method of action recognition, detection and clustering by analyzing concatenated 2-d shapes in a video time sequence.

The method use poisson equation to extract shape properties of human actions. Extracting shape silhouette is achieved by with fixed camera position. Thus background recognition is simple. Poisson equation characterize the shape space-time points with local orientation and aspect ration of shape. And those local properties and parameters are integrated into global features to represent a action.

The moving part and body torso are identified by computing the gradient over x, y and t. A space-time saliency feature is used by normalizing the gradient and gives more preference to fast moving part. Space orientations are define by the principle eigenvectors of Hessian of passion equation. Global feature is computed by integrating the saliency features and space orientation.

The method is straightforward. Action is classified by NN with regardless of person labels. The algorithm misclassified 20 out 923. It's a high accuracy, however the dataset is clean and background is distinctive from human object. But the algorithm is sensitive to the mesh ratios of time and shape.Those parameters are chosen by empirical experiments. The space-time cubes are overlapped and generated by sliding window in time. One reason might be the duration of human actions is uncertain. Using overlapped cubes avoids the partial occlusion and noise.