

Motion segmentation using GPCA techniques and optical flow Segmentación de movimiento utilizando técnicas GPCA y flujo óptico

C. Losada, M. Mazo, S. Palazuelos, J. L. Martín, J. J. García
Departamento de Electrónica. Universidad de Alcalá
Alcalá de Henares, Madrid, España
losada@depeca.uah.es

Abstract

In this work, the use of the Generalized Principal Components Analysis (G-PCA) to improve the segmentation of moving objects in image sequences is proposed. In order to obtain this improvement, the noise components in the image derivatives are reduced, and afterwards, a method based on linear algebra is used to make the segmentation. Furthermore this work presents diverse tests to compare the results reached with and without the noise reduction in the image derivatives, and using the nonlinear minimization of an error function. A remarkable improvement in the segmentation quality and the processing time can be observed in every experiment when using the proposed method.

Keywords: GPCA, motion segmentation, optical flow.

Resumen

En este trabajo se propone la utilización del Análisis de Componentes Principales Generalizado (G-PCA) para mejorar la segmentación de objetos movimiento en secuencias de imágenes. Esta mejora se consigue tras realizar una reducción de las componentes de ruido en las derivadas de las imágenes, que luego se utilizarán para la segmentación mediante un método basado en el álgebra lineal. Se han realizado diversas pruebas en las que se comparan los resultados alcanzados al realizar la reducción del ruido en las derivadas de la imagen, con los que se obtienen empleando únicamente el algoritmo algebraico de segmentación, e incluso con los proporcionados mediante minimización no lineal de una función de error. En todos los casos, se ha observado una notable mejoría con el método propuesto en este trabajo.

Keywords: GPCA, segmentación de movimiento, flujo óptico.

1. Introducción

La segmentación de objetos en movimiento es una tarea fundamental para el análisis de secuencias de

imágenes y consiste en agrupar la información presente en las mismas en conjuntos de píxeles cuyo movimiento en el plano imagen es coherente a lo largo de una secuencia. Todo ello sin tener un conocimiento previo acerca de qué píxeles de la imagen se mueven de acuerdo a un determinado modelo de movimiento. El análisis de movimiento en secuencias de imágenes digitales tiene un gran interés para una amplia gama de aplicaciones de la visión artificial, tales como robótica móvil, monitorización de tráfico, registrado de imágenes, vigilancia, etc.

Existen en la literatura diferentes técnicas que tratan de analizar y extraer información sobre el movimiento de los objetos en una escena. Las primeras aproximaciones para segmentación de movimientos 2D son las basadas en discontinuidades en el flujo óptico [1], [2], pero estas técnicas presentan diferentes inconvenientes derivados tanto del problema de apertura, como de la presencia de ruido en las estimaciones del flujo óptico.

Por otro lado, la segmentación del movimiento ha estado tradicionalmente acoplada con la detección de movimiento, en la que cada región corresponde a un modelo particular de movimiento que explica los cambios temporales en dicha región de la imagen [3], [4]. También se han realizado diferentes propuestas para segmentación de movimiento basadas en clustering [5], [6], [7], sin embargo, en cualquiera de los casos anteriores es necesario tener un conocimiento a priori del número de objetos en movimiento presentes en la escena, lo cual no siempre es posible.

En [8] y [9] se propone un algoritmo de segmentación de objetos con modelos de movimiento aún basado en la técnica matemática GPCA (Generalized Principal Components Analysis) que permite realizar la segmentación de los objetos en movimiento sin tener un conocimiento previo del número de objetos presentes en la escena, y evitando el problema de apertura en el cálculo del flujo óptico.

La estructura de este trabajo es la siguiente: en la sección 2 se describen muy brevemente los algoritmos de [8] y [9]. En el apartado 3 se presenta la mejora propuesta. A continuación, en la sección 4, se muestran y comparan los resultados obtenidos con los diferentes

algoritmos, y por último, en la sección 5 se exponen las conclusiones extraídas de este trabajo.

2. Segmentación de movimientos afines 2D utilizando la técnica GPCA

Dado un conjunto de N medidas $\{(\mathbf{x}_j, \mathbf{y}_j)\}_{j=1}^N$ de imágenes, correspondientes a un número desconocido de movimientos afines, es posible utilizar la técnica matemática GPCA para estimar el número de movimientos n , los parámetros de cada movimiento afín $\{\mathbf{A}_i\}_{i=1}^n$ y realizar la segmentación de las imágenes medidas en función de los modelos de movimiento estimados, siendo $\mathbf{y} = [I_{x_1} \ I_{x_2} \ I_t]^T \in \mathfrak{R}^3$ un vector que contiene las derivadas espaciales y temporales de la imagen, y $\mathbf{x} \in \mathbb{P}^2$ las coordenadas homogéneas de los píxeles.

Cualquier medida en la imagen (\mathbf{x}, \mathbf{y}) asociada con algún movimiento debe satisfacer la restricción afín multicuerpo definida en la ecuación (1). Esta restricción reduce el problema de segmentación a obtener el número de movimientos afines n y los parámetros de dichos movimientos a partir de la ecuación no lineal:

$$\varepsilon(\mathbf{x}, \mathbf{y}) = \prod_{i=1}^n (\mathbf{y}^T \mathbf{A}_i \mathbf{x}) = 0 \quad (1)$$

Dada la estructura de esta ecuación, es posible utilizar la inmersión de Veronese de grado n , $v_n : \mathfrak{R}^3 \rightarrow \mathfrak{R}^{M_n}$ para reescribir la restricción afín multicuerpo en forma bilineal:

$$v_n(\mathbf{y})^T \mathbf{A} v_n(\mathbf{x}) = 0 \quad (2)$$

donde $\mathbf{A} \in \mathfrak{R}^{M_n \times M_n}$ es la matriz afín multicuerpo que representa el producto tensorial simétrico de todas las matrices afines $\{\mathbf{A}_i\}_{i=1}^n$. La restricción anterior puede describirse:

$$(v_n(\mathbf{y}) \otimes v_n(\mathbf{x}))^T \mathbf{a} = \mathbf{L}_n \mathbf{a} = 0 \quad (3)$$

donde \otimes indica producto tensorial. La expresión (3) nos permite calcular el número de movimientos n imponiendo una restricción sobre el rango de la matriz \mathbf{L}_n de forma que el sistema de ecuaciones (3) tenga una solución única. Conocido el valor de n , la matriz afín multicuerpo \mathbf{A} se obtiene resolviendo el sistema de ecuaciones lineales en \mathbf{a} .

Como se demuestra en [8], GPCA nos permite recuperar el flujo óptico $\{\mathbf{u}_i\}_{i=1}^n$ a partir del flujo óptico multicuerpo $\tilde{\mathbf{u}} = \mathbf{A} v_n(\mathbf{x})$ de forma sencilla con cualquiera de los algoritmos para hiperplanos propuestos en [9] y

posteriormente recuperar los modelos de movimiento afín individuales, así como realizar la segmentación de movimiento.

Como ya se ha comentado, esta propuesta para segmentación de movimiento tiene algunas ventajas ya que no requiere un conocimiento previo del número de modelos de movimiento presentes en la escena (proporciona un método para calcularlo) y evita el problema de apertura en el cálculo del flujo óptico. Sin embargo, aunque el problema es no lineal, el algoritmo propuesto está basado en técnicas del álgebra lineal tras realizar el embebido de los datos mediante la inmersión de Veronese. Esto provoca que, a pesar de que la solución propuesta es algebraicamente correcta, la estimación de la matriz afín multicuerpo no es robusta en presencia de ruido en las medidas de la imagen.

En [8] se propone tratar las medidas ruidosas realizando un proceso de optimización que minimice una función del error algebraico definida por la restricción afín multicuerpo, mediante técnicas de minimización no lineal, utilizando el algoritmo algebraico para realizar la inicialización.

Sin embargo, dada la naturaleza del problema, a medida que el ruido aumenta, el tiempo de convergencia del algoritmo también aumenta. Por otro lado, también es posible que la minimización devuelva una solución local, en lugar de la solución global del problema. Por esto no resulta adecuado para realizar la segmentación de forma robusta, especialmente en casos en que existen restricciones de tiempo.

3. Mejora de la segmentación de movimiento mediante la reducción del ruido en las derivadas de la imagen

A la vista de los problemas que presenta la solución anterior, derivados de la presencia de ruido en las derivadas (\mathbf{I}_{x_1} , \mathbf{I}_{x_2} e \mathbf{I}_t) obtenidas de las imágenes captadas (que a partir de ahora denominaremos “imágenes de movimiento”) se propone aplicar una variante del algoritmo algebraico propuesto en [8], pero tratando de eliminar previamente las componentes de \mathbf{I}_{x_1} , \mathbf{I}_{x_2} e \mathbf{I}_t que sean poco significativas y que, en consecuencia, puedan ser debidas a la presencia de ruido.

Para la reducción del efecto del ruido, se emplea el análisis de componentes principales (PCA) aunque, dada la estructura matricial de las derivadas de la imagen, en lugar del algoritmo PCA clásico [10] para la obtención de las matrices de transformación, se ha decidido utilizar el algoritmo PCA Generalizado (de ahora en adelante G-PCA) propuesto en [11], que, a pesar de la coincidencia en el nombre, no guarda relación con la técnica GPCA de [8] y [9] para segmentación.

3.1. Obtención de las componentes de flujo óptico significativas en las imágenes de movimiento

La propuesta realizada para la reducción del efecto de ruido en las componentes del flujo óptico (derivadas espaciales y temporales) se presenta de forma esquemática en la Figura 1. Como se puede observar, se parte de r imágenes de una escena tomadas por una única

cámara. De cada par de imágenes (Im_i, Im_{i+1}) se obtienen las derivadas espaciales I_{x1i} e I_{x2i} y la derivada temporal I_t . A partir de estos datos es posible obtener las matrices de transformación $R_{x1}, L_{x1}, R_{x2}, L_{x2}, R_t$ y L_t utilizando el algoritmo G-PCA [11]. Más adelante se explicará la técnica empleada para calcular el número de componentes principales d necesarias en cada caso.

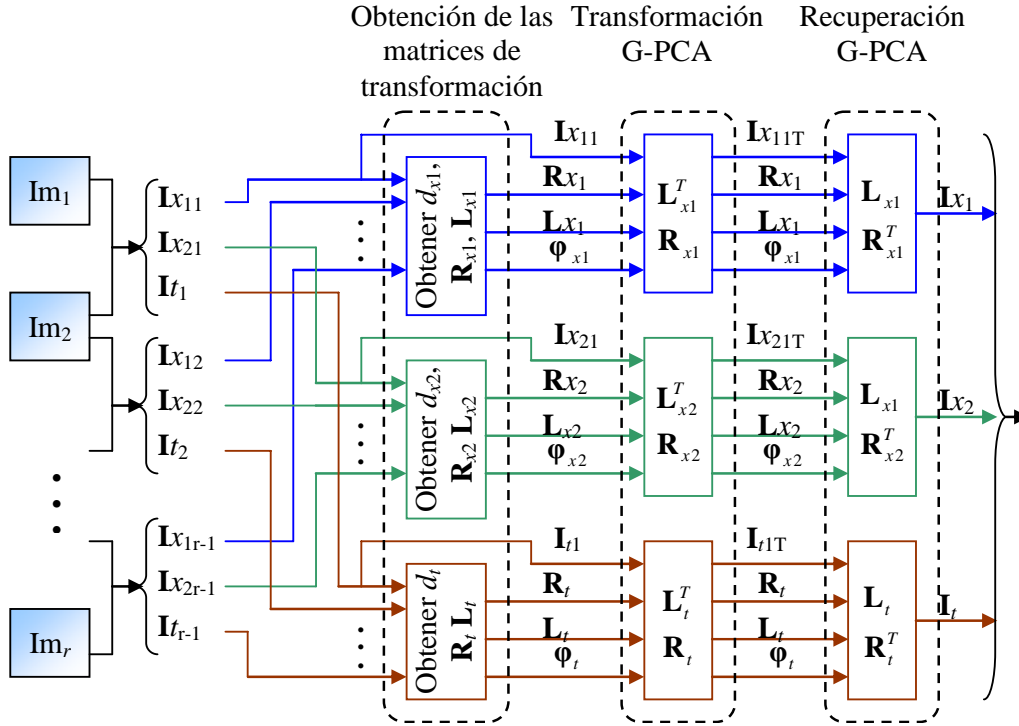


Figura 1. Diagrama de bloques propuesto para la reducción de las componentes de ruido en las derivadas de las imágenes

Una vez se han obtenido las matrices de transformación es posible proyectar los datos en el espacio transformado (4), (5) y (6), siendo ϕ_{x1} , ϕ_{x2} y ϕ_t las medias de las matrices I_{x1} , I_{x2} e I_t respectivamente.

$$I_{x1T} = L_{x1}^T (I_{x11} - \phi_{x1}) R_{x1} \quad (4)$$

$$I_{x2T} = L_{x2}^T (I_{x21} - \phi_{x2}) R_{x2} \quad (5)$$

$$I_{tT} = L_t^T (I_{t1} - \phi_t) R_t \quad (6)$$

Posteriormente se recuperan los datos mediante la transformación inversa (7), (8), (9) de forma que, en las matrices recuperadas, se conserve la información de los objetos en movimiento presentes en la escena, pero se hayan reducido las componentes de ruido.

$$I_{x1R} = L_{x1} I_{x11T} R_{x1}^T + \phi_{x1} \quad (7)$$

$$I_{x2R} = L_{x2} I_{x21T} R_{x2}^T + \phi_{x2} \quad (8)$$

$$I_{tR} = L_t I_{t1T} R_t^T + \phi_t \quad (9)$$

3.2. Obtención del número de componentes principales d

El buen funcionamiento de esta técnica para la eliminación de ruido depende, en gran medida, de la correcta elección del número de componentes principales elegido d . Con objeto de eliminar la mayor cantidad de ruido en las imágenes de movimiento, sin perder la información que contienen dichas imágenes, se ha realizado un estudio sobre los autovalores de las matrices de transformación. Para ello se aplica una única iteración del algoritmo de [11] y se elige el valor de d como el número de autovalores de la matriz R cuyo valor supera el 10% del autovalor máximo.

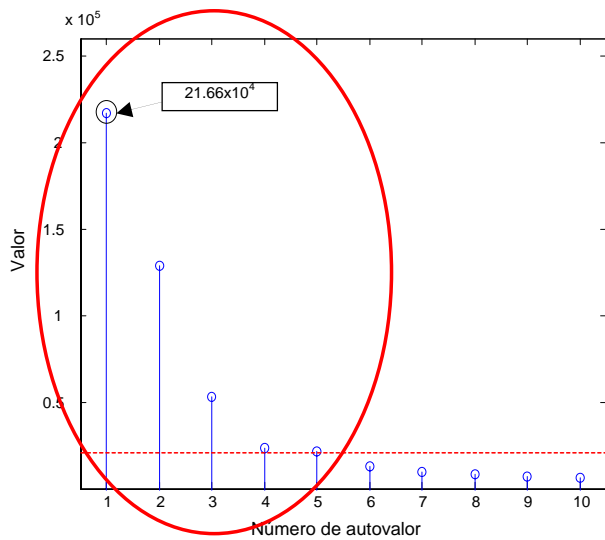


Figura 2. Ejemplo de autovalores de la matriz R para I_{x1} obtenidos a partir de r = 4 imágenes

En la Figura 2 se presenta un ejemplo en el que se seleccionarían los 5 componentes ya que, como se puede observar, son los que superan el 10% del valor del autovalor máximo.

3.3. Segmentación de objetos en movimiento a partir de las derivadas de la imagen

Para la segmentación de movimiento se emplea el algoritmo algebraico propuesto en [8], cuyas etapas se muestran en el diagrama de bloques de la figura 3. Los datos de entrada a este algoritmo son las matrices obtenidas tras la transformación y posterior recuperación de las imágenes de movimiento I_{x1}, I_{x2} e I_t.

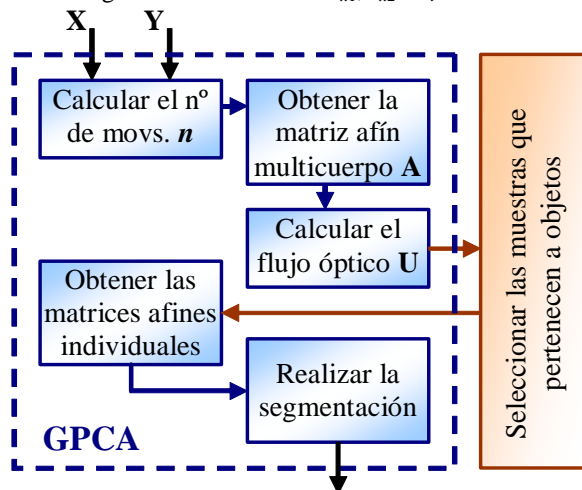


Figura 3. Diagrama de bloques para segmentación de movimiento propuesto

La selección de las N muestras $\mathbf{X} = [x_1^j \ x_2^j \ 1]^T_{j=1..N}$ e $\mathbf{Y} = [I_{x1}^j \ I_{x2}^j \ I_t^j]^T_{j=1..N}$ que pertenecen a objetos en movimiento presentes en la imagen se realiza aplicando la restricción afín multicuerpo (2) tras el cálculo de la matriz afín multicuerpo de la forma que se indica en [8], e imponiendo un umbral al resultado, de forma que, únicamente las muestras que se encuentran por debajo de ese umbral se consideran en las siguientes etapas del algoritmo. Así se produce una reducción del tiempo de cómputo necesario para la ejecución del algoritmo.

4. Resultados

Para comprobar las prestaciones del algoritmo propuesto se han comparado los resultados que proporciona esta solución con los obtenidos mediante el algoritmo [8]. En nuestro caso el objeto a segmentar es un robot móvil.

Todas las pruebas que se presentan aquí han sido realizadas utilizando 4 imágenes para obtener las matrices de transformación G-PCA, aunque también se han obtenido resultados muy satisfactorios empleando únicamente 3 imágenes con este fin. En cuanto a la segmentación del robot, en la Figura 5 se presentan los resultados obtenidos al aplicar las diferentes alternativas presentadas para segmentación de movimiento sobre cada par de imágenes, de forma que, por cada 4 imágenes de entrada, obtendremos tres imágenes segmentadas.

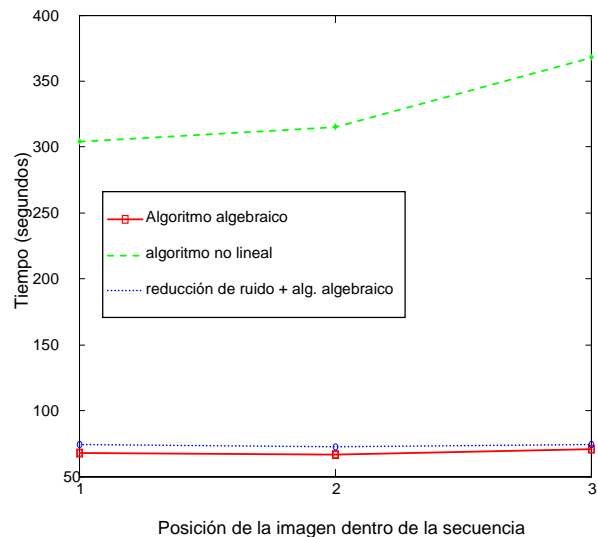


Figura 4. Comparación de los tiempos de cómputo del algoritmo propuesto con el presentado en [8]

Se puede observar como con la propuesta de reducción de ruido, previa a la segmentación, se reduce la detección de falsos puntos, esto es puntos estáticos que se

consideran con el mismo movimiento que el robot (puntos detectados que están fuera del robot).

También se han comparado los tiempos de ejecución del algoritmo propuesto con el presentado en [8]. En la Figura 4 se representan los tiempos de ejecución de cada uno de estos algoritmos para la segmentación de cada par de imágenes de prueba.

Se puede observar cómo la introducción de una nueva etapa para la reducción del ruido produce un aumento en el tiempo de ejecución con respecto al caso en que se emplea únicamente el algoritmo algebraico para minimización. Sin embargo, este tiempo es siempre mucho menor que el consumido por el algoritmo de minimización no lineal.

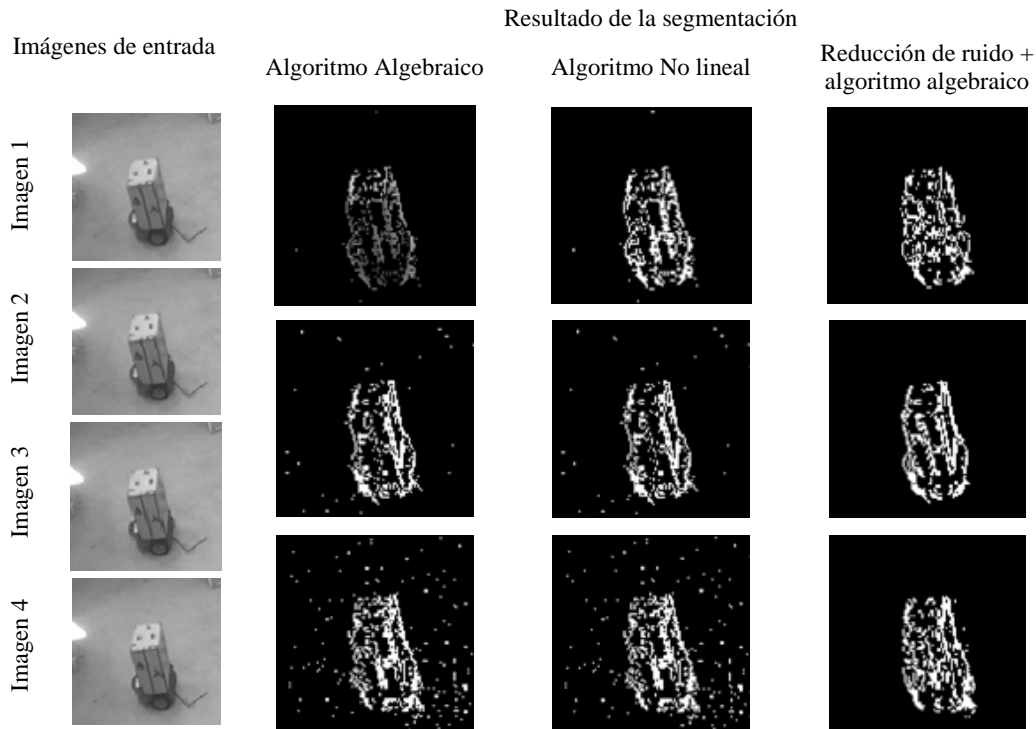


Figura 5. Ejemplos de resultados de segmentación obtenidos aplicando diferentes alternativas.

5. Conclusiones

Actualmente existen diferentes alternativas que permiten realizar la segmentación de los modelos de movimiento presentes en una escena, pero todas ellas presentan problemas que no las hacen adecuadas en los casos en que se requiera una segmentación robusta.

La técnica matemática GPCA permite realizar la segmentación de movimiento sin necesidad de conocer de forma previa el número de objetos en movimiento presentes en la escena, sin embargo, su comportamiento se degrada rápidamente en presencia de ruido en las derivadas de las imágenes. Para tratar de disminuir el efecto del ruido, es posible resolver el problema de segmentación aplicando técnicas de minimización no lineal sobre una función de error definida por la restricción afín multicuerpo, pero dada la naturaleza del

problema, los resultados obtenidos por este método tampoco son completamente satisfactorios, con la desventaja añadida del aumento en el tiempo de proceso.

En este trabajo se ha propuesto una alternativa que permite la segmentación de uno o varios objetos en movimiento presentes en una secuencia de imágenes tomadas por una cámara, introduciendo una etapa previa a la segmentación en la que, mediante un técnica de Análisis de Componentes Principales Generalizado (G-PCA) se reducen las componentes de ruido en las derivadas de la imagen.

En las pruebas realizadas se ha puesto de manifiesto que la reducción del ruido en las derivadas de las imágenes de forma previa a la segmentación mejora notablemente los resultados obtenidos. Además, aunque el tiempo de proceso es algo mayor que en el caso de emplear únicamente el algoritmo algebraico, se ha

conseguido una aceleración significativa con respecto a la utilización del algoritmo de minimización no lineal.

Agradecimientos

Este trabajo ha sido posible gracias a la financiación del Ministerio de Educación y Ciencia (MEC) a través del proyecto RESELAI (REF-TIN2006-14896-C02-01).

Referencias

- [1] Black, M. and Anandan, P. "Robust dynamic motion estimation over time". Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition. June 1991. Page(s): 296-302.
- [2] Yan, H.; Tjahjadi, T.; "Multiple motion segmentation through a highly robust estimator". IEEE International Conference on Systems, Man and Cybernetics Oct. 2004. Page(s): 3082-3087.
- [3] Darrell, T.; Pentland, A. "Robust estimation of a multi-layered motion representation". Proc. of the IEEE Workshop on Visual Motion. 1991. Page(s): 173-178.
- [4] Weiss, Y. "A unified mixture framework for motion segmentation: incorporating spatial coherence and estimating the number of models". Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition. 1996. Page(s): 321—326
- [5] Boulton, T.E. and Brown, L.G. "Factorization-based segmentation of motions". Proc. of the IEEE Workshop on Motion Understanding. pages 179–186 (1991)
- [6] Costeira, J. and Kanade, T. "A multibody factorization method for independently moving objects". International Journal of Computer Vision. Volume 29, Number 3. September 1998. Page(s):159–179.
- [7] Kanatani, K. "Motion Segmentation by Subspace Separation and Model Selection". Proceedings of the 8th IEEE International Conference on Computer Vision 2001. Volume 2, Page(s): 586-591
- [8] R. Vidal and S. Sastry. "Segmentation of Dynamic Scenes from Image Intensities". IEEE Workshop on Vision and Motion Computing. Page(s): 44-49. 2002.
- [9] Vidal, R.; Yi Ma; Sastry, S.; "Generalized principal component analysis (GPCA)". IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 27, Issue 12, Dec. 2005 Page(s):1945 – 1959
- [10] Turk, M., Pentland, A., "Eigenfaces for recognition", Journal of Cognitive Neuroscience, Vol.3, pp.72-85 (1991).
- [11] Ye, Jieping; Janardan, Ravi; Li, Qi. "GPCA: an efficient dimension reduction scheme for image compression and retrieval". Proceedings of the 10th ACM SIGKDD international conference on Knowledge discovery and data mining. 2004. Pages: 354 – 363.