Technical Research Paper:

Object Tracking

Kevin Andrade

Binghamton University

**Thesis**

The partial least squares algorithm (PLS), when compared to other algorithms such as the multiple instance learning algorithm (MIL) does a much better job at tracking objects in difficult visual scenarios, while improving the effectiveness of security and surveillance cameras, human-phone interaction, medical imaging, and traffic control.

**Introduction**

Object tracking software is on its way to being the new set of eyes for technology. Object tracking software can track and follow an object with its complex algorithms. Object tracking is implemented in various fields, and is used to improve the efficiency of human computing technology. There are many algorithms available in the field of object tracking, but some are better than others. The partial least squares algorithm (PLS), when compared to other algorithms such as the multiple instance learning algorithm (MIL) does a much better job at tracking objects in difficult visual scenarios, while improving the effectiveness of security and surveillance cameras, human-phone interaction, medical imaging, and traffic control. Tracking objects in a video depends on several factors, such as the quantity and shape of the object, and the visual scenario that the object is placed in. Philip et al. (2014) compares different algorithms against each other through a variety of experiments. Wang et al. (2012) further explains the handling of object analysis and tracking via the PLS algorithm; the authors also describe the issues of the PLS algorithm and its performance through a variety of experiments with different types of visual scenarios. Object tracking algorithms are a very popular research topic.

**Background**

Object tracking is a widely-studied topic, and is used in many fields, such as: the surveillance field, the medical field, and the robotics field (Wang, 2012). Object tracking gives an application the ability to track an object using complex algorithms; while object tracking has developed paths to improve different fields, there are some tracking algorithms that are superior to others. When in use, tracking algorithms depend on several factors, such as: the visual scenario, the prior knowledge about the object, and the number and shape of the object (Philip, 2014). Normally, a tracking system is made from three parts: the evaluation model, which tries to evaluate where the object is, the appearance model, which relates the locations of objects over time, and the search strategy, which finds the most likely location of the current image (Philip, 2014). Tracking objects in high resolution video has been successful in the last 2 decades, but tracking objects in a low-resolution video has been challenging (Philip, 2014). If an object moves fast relative to the frame rate of the video, then the image model is lost. Another complexity is when an object changes orientation over time, it makes it hard to track the dimensional shape of the object.

The ability to track human faces via surveillance cameras is essential to classify criminals. There are many criminals who escape crimes via cars, and it is difficult to track fast movement due to the occlusion handling, the motion blur, the change of appearance, and the change in illumination. Object tracking software is also used in military planes to track and lock onto enemies. Being able to analyze the human body via tracking algorithms is essential to the medical field; there are cases when a person must undergo surgery, and it is always a good idea to know the condition of the body before going into a medical procedure (Metaxas, 2005). The ability to track objects has revolutionized the industry, with the partial least squares algorithm being one of the superior algorithms.

**Precedents and related work**

Before object tracking algorithms were robust, there were different methods to track an object; people used software to go through a video frame by frame to track motion, this was long and tedious work. Object tracking has improved greatly over the recent years, with researchers and scientist creating and improving previous algorithms. There are numerous algorithms that will track objects at the click of a button versus the long and tedious way of going frame by frame to localize and track an object manually. Existing algorithms are categorized as either generative or discriminative methods (Wang, 2012). Generative methods track objects by searching the objects image and matching it with a template or appearance model in their database. Discriminative methods try to separate the object from its background. Despite the ease of newer tracking technologies, it isn’t always perfect. Real world object tracking is difficult due to the many visual changes happening throughout the tracking cycle. Sometimes, the object being tracked is partially hidden behind another object, and this causes a problem. Philip et al. (2014) compared multiple object tracking algorithms by putting them through various tests. He concluded that the PLS algorithm was overall the most accurate tracking algorithm when compared to other algorithms such as the MIL algorithm.

When dealing with occlusion handling, negligible motion, background clutter, and low contrast video, the PLS algorithm does much better job when compared to the MIL algorithm; the difference in the succession rate address the many social issues of inadequate object tracking.

There are hospitals that specialize in motion analysis, and every day, the need for stronger tracking software is desired (Connecticut Children’s office, 2017). The PLS algorithm aids the development of strong tracking software, which results in the improvement of medical practices. Hospitals can now track the posture and movement of joints with their video tracking software (Connecticut Children’s office, 2017), and are able to track internal organs to determine the patients’ health status (Metaxas, 2005), but the algorithms can still be improved. The results from Philip et al.’s tests concluded that the PLS algorithm was overall the best algorithm, but it still had complications keeping track of objects 100% of the time due to the difficult visual scenarios.

Real life cases such as traffic control, human and robot interaction, and medical imaging rely heavily on object tracking algorithms to track an object efficiently. War criminals can be tracked and identified via smart surveillance cameras which match their face via an online database, complicated organs can be studied and examined without having to go through surgery, and the government can have a more reliable system of tracking enemies from an aerial view via the PLS algorithm.

**Object Representation**

Wang et al. (2012) describes how the PLS algorithm works and on how the PLS algorithm can analyze and track images, even when the object being tracked is in a difficult visual scenario with superb tracking capabilities. Wang et al. states that the PLS algorithm is a statistical method for modeling the relations between sets of variables via latent quantities. Latent variables are variables that are not observed, but are rather inferred. Object tracking is a classification problem with the PLS algorithm to learn low dimensional and discriminative feature subspace.

1. **Partial Least Squares Analysis**

Wang et al. (2012) uses different types of dimensional space to describe the PLS analyses. Wang et al. let’s X ∈ R^m be an m dimensional space of variables and Y∈ R^m be an n dimensional space of other variables. With a quantity of N observed samples from x ∈ X and y ∈ Y that form two distinct blocks of variables (X ∈ R^(Nxm) and Y ∈ R^(Nxm). The PLS method finds spaces and then preserves the observed samples, and the learned latent variables from the two blocks are more correlated than those in the original spaces. (Equation 1 X = TP’ + E, Y = UQ^T + F) where T ∈ R ^(Nxp) and U ∈ R^(Nxp) are factor matrices, P ∈ R^(mxp) and Q ∈ R^(nxp) are loading matrices, and E ∈ R^(Nxm) and F ∈ R^(Nxn) are error terms. Each variable is represented by a p dimensional vector. To solve for X and Y in equation 1, the pls algorithms first compute the weight vectors w1 and c1, where most of the variations in X and Y are retained by t1 = Xw1 and u1 = Yc1, (Equation 2 max w1 V ar(t1), max c1 Var(u1) where t1 and u1 are the first columns of T and U, and V ar(.) is the variance. T1 has to to explain u1 (Equation 3 max w1, c1 ρ(t1, u1) ) where p(t1, u1) = Cov(t1, u1)/√V ar(t1)Var(u1) is the connection and correlation between t1 and u1, and Cov(t1, u1) = t ^T1 u1/N defines the sample correlation between t1 and u1. By combining equations 2 and 3, the analysis and correlation of the pls algorithm is maximized between t1 and u1 (Equation 4 ) from this equation we can solve the following problem (Equation 5 max (Xw1, Y c1) s.t. w1 w1 = 1, c1 c1 = 1 ) where (Xw1, Yc1) indicate the product of Xw1 and Yc1, with w1 being the first eigenvector of the proceeding eigenvalue problem. (Equation 6 XY^(T) Y^(T) Xw1) = λw1). c1 can be obtained by a solving another problem by a similar manner. (Equation 7) Y^(T) XX ^(T)Yc1 = = λc1. After performing the first step, the PLS method iteratively defines other weight vectors. After solving for w1 and c1 the score vectors can be computed by using t1 = Xw1, u1 = Yc1, and loadings can be computed with the equation p1 = (x^(T)t1)/(t1^(T)t1) and q1 =(Y^(T)u1) / (u1^(T)u1). X and Y are subtracted by their rank one approximations (Equation 8 X <- X – t1P1, Y <- Y –u1q1). With new values for X and Y, it is now possible to compute w2, c2 using equation 6 and 7. The process gets repeated until the residuals become a small quantity of weight vectors. Low dimensional space is learned when latent quantities from different sets of observed variables are more correlated when compared to those of original spaces. PLS has become important in many fields, with the help of tracking, spotting people has become a lot more efficient.

**Experiments**

Philip et al. (2014) conducted a few experiments with a variety of cases with difficult visual scenarios and compared them to other algorithms. Object tracking is difficult due to the many inconsistencies of video, but the PLS system seems to perform the best overall when compared to other algorithms such as the mil algorithm. The experiments conducted were shot from an aerial view with a video resolution of 720 x 480p, with approximately 39 cars going across all 50 frames.

1. **Overall Accuracy**

Philip et al. measured accuracy based on two metrics 1) localization error (the error in localizing the correct object) 2) overlap accuracy (when another object gets in front of targeted object). As show in the figure, the PLS algorithm performed the best overall when compared to other algorithms. Other algorithms are comparable in their overlap accuracy percentage, excluding the MIL algorithm being last in the list

(Philip et al., 2014)

1. **Occlusion Handling**

While the test was being conducted, car number 36 was a big challenge for the tracking algorithms. Between frames 18-27 car number 36 was partially hidden behind a traffic light which

As mentioned previously, object tracking is based on classification by labeling the target positive and the background negative. Wang et al. (2012) described the mathematical formulas on how the PLS algorithm analysis. Wang et al. uses X ∈ R^m to denote feature space for object description, and Y ∈ R to describe the class label space of an object. Once the object is tracked, we have a positive sample x1 by extracting a feature vector from the warped image specified by the state parameter.

.

Support outline

**References**

Philip, R. C., Ram, S., Gao, X., & Rodriguez, J. J. (2014, April). *A comparison of tracking algorithm performance for objects in wide area imagery.* 2014 Southwest Symposium on Image Analysis and Interpretation, San Diego, California, USA. doi:[10.1109/SSIAI.2014.6806041](https://doi.org/10.1109/SSIAI.2014.6806041)

Wang, Q., Chen, F., Xu, W., & Yang, M. (June 2012). Object tracking via partial least squares analysis*. IEEE Transactions on Image Processin*g, *21*, 4454-4465. doi: 10.1109/TIP.2012.2205700

Metaxas, D. N. (February 2005). Introduction. *Communications of the ACM*, *48*, 26-29. doi: [10.1145/1042091.1042115](http://dx.doi.org.proxy.binghamton.edu/10.1145/1042091.1042115)

Connecticut Children’s Office. (2017). About motion analysis. Retrieved from https://www.connecticutchildrens.org/orthopedics-center-for-motion-analysis/what-is- motion-analysis/