# Multiple Sensors Based Hand Motion Recognition Using Adaptive Directed Acyclic Graph

**Yaxu Xue [1,2,3], Zhaojie Ju [1,2,*], Kui Xiang [1], Jing Chen [1] and Honghai Liu [2]**

[1]  School of Automation, Wuhan University of Technology, Wuhan 430070, China;
    yaxu.xue@whut.edu.cn (Y.X.); xkarcher@126.com (K.X.); jingchen680@163.com (J.C.)
[2]  School of Computing, The University of Portsmouth, Portsmouth PO1 3HE, UK; honghai.liu@port.ac.uk
[3]  School of Electrical and Mechanical Engineering, Pingdingshan University, Pingdingshan 467000, China
[*]  Correspondence: zhaojie.ju@port.ac.uk; Tel.: +44(0)23-9284-5382

**Abstract:** The use of human hand motions as an effective way to interact with computers/robots, robot manipulation learning and prosthetic hand control is being researched in-depth. This paper proposes a novel and effective multiple sensor based hand motion capture and recognition system. Ten common predefined object grasp and manipulation tasks demonstrated by different subjects are recorded from both the human hand and object points of view. Three types of sensors, including electromyography, data glove and FingerTPS are applied to simultaneously capture the EMG signals, the finger angle trajectories, and the contact force. Recognising different grasp and manipulation tasks based on the combined signals is investigated by using an adaptive directed acyclic graph algorithm, and results of comparative experiments show the proposed system with a higher recognition rate compared with individual sensing technology, as well as other algorithms. The proposed framework contains abundant information from multimodal human hand motions with the multiple sensor techniques, and it is potentially applicable to applications in prosthetic hand control and artificial systems performing autonomous dexterous manipulation.

**Keywords:** EMG; contact force; data glove; adaptive directed acyclic graph

## 1. Introduction

As an extraordinarily dexterous part of the human body, the human hand achieves most of the tasks in daily life. Hand movements provide a natural and intuitive manipulation modality while performing behaviors including gestures, postures, and manipulations. With the high-speed development of modern robot technology, robots are being used to work in more and more complex surroundings like aerospace, field operations, and repeated or dangerous works [1]. This technology has a profound impact on traditional industries, but faces some challenging problems. These applications require autonomous dexterous robots that perform increasingly human-like manipulation tasks and achieve advanced manipulation tasks such as in-hand regrasping, rotation and translation. The current level of development of a sophisticated multi-fingered robot hand is still at an early stage because of its control complexity and the immature synchronous cooperation between sensor-motor systems [2]. Hence, how to intensively investigate in-hand manipulation skills of humans, and further transfer such skills into bionic multi-fingered dexterous robotic hands have been receiving considerable attention [3].

In order to study the effectiveness of the dexterous human hand manipulation, it is crucial to accurately extract its multi-modal information, such as fingertip locations, hand skeleton, contact force, speed, etc. [4]. Current hand motion capturing systems can be mainly categorised into: data glove based capturing, attached force based capturing, electromyography (EMG) based capturing, optical markers

based capturing and vision based capturing. Increasing efforts have been made in analysing the hand motions by using uni-modal sensors. Luzanin et al. developed a data glove-based hand motion recognition system which used a probabilistic neural network trained on a cluster set generated by a clustering ensemble [5]. Some mainstream sensors, including capacitive, piezoresistive, piezoelectric and optoelectronic techniques, were designed to detect 3D contact force and physical properties of object [6]. By utilizing multichannel surface EMG, individual and combined finger movements were classified for dexterous prosthetic control, and then offline processing was used to evaluate the classification performance [7]. Metcalf et al. presented a kinematic model based on surface marker placement and used standard calculations to calculate the movements of the wrist, hand, fingers and thumb [8]. A comprehensive review of recent Kinect-based computer vision algorithms and applications, including preprocessing, object tracking and recognition, human activity analysis, hand gesture analysis, and indoor 3D mapping is presented in [9]. In addition to the literature mentioned above, a large number of related scientific papers as well as technical demonstrations have already appeared in various robot journals and conferences. However, most of the current research focused on a range of limited behaviours or in limited scenarios.

Because of the complex properties involved in the human hand motions, hand motion capture will be faced with some additional challenges, like large posture variations, different colored skin and severe occlusions of the fingers during movements. A single sensor can get a certain feature information for some simple gesture recognition, such as static gestures. However, due to its own limitations, the lack of some key features may cause movement distortion in the complex gesture recognition, such as in-hand manipulation and re-grasps. It is crucial to compensate the drawbacks of uni-modal sensors by using a combination of varied types of sensors. For instance, data gloves do not have the basic ability of the spatial location information collection, and thus vision based sensors like Kinect, Leap motion controller can be integrated into motion capture system to compensate the limitation of data glove [10]. Hence, given the difficulty of robust control based on uni-modal sensors, multimodal sensory fusion is a promising approach. Some researchers have done great work on the integration of multiple types of sensors for analyzing the human hand motions. Park et al. employed an artificial skin with both of the multi-axis strains and contact pressure information to analyze the human hand motions [11]. Multimodal tactile sensor (BioTac) was applied to measure overall internal fluid pressure and electrode impedance for haptic perception of the small, finger-sized geometric features [12]. A generalized framework integrating multiple sensors to study and analyze the hand motions that contained multimodal information was proposed in [13]. Based on the vision sensors, Marin et al. proposed a novel hand gesture recognition scheme explicitly targeted at the hand motion analysis [14]. More examples could be found in [15], which presented a comprehensive review of the current state of the biosensing technologies focusing on the hand motion capturing and their applications on hand prostheses. However, most references show that multiple sensory information is integrated through two types of sensors. There are few researchers to analyze the hand motions by using the integration of muscle signals with the finger trajectories and the contact forces.

This paper designs a multiple sensor based hand motion capture system, which includes EMG, data glove, and Finger Tactile Pressure Sensing (FingerTPS) for hand motion recognition. The rest of the paper is organized as follows. First, Section 2 describes the current multiclass support vector machines (MSVMs) methods and analyses their advantages and disadvantages, and then the Adaptive Directed Acyclic Graph (ADAG) algorithm is applied to recognise different hand motions based on multiple sensory information. Then, the proposed hand motion capture system and the related preprocessing module are introduced in Section 3. In addition, an improved grid-search approach is used to get the optimal intrinsic parameters. The capability of the proposed hand motion recognition algorithm and the comparative experimental results have been investigated in Section 4. Finally, Section 5 concludes this paper with further discussion.

## 2. Hand Motion Recognition Method

The hand motion recognition module is to identify both different hand grasps and in-hand manipulations. In this section, we will firstly introduce the recognition method and then discuss the comparative experimental results.

### 2.1. Multiclass Support Vector Machines

Support Vector Machine (SVM) is a novel large margin classifiers used for classification and regression. SVM maps vectors onto a much higher dimensional space, and sets up a maximum margin hyperplane that divides clusters of vectors. New objects are then located in the same space and were recognized to belong to a class based on which area they fall in. This method can minimize the upper bound of generalization error and provide excellent generalization ability, so it is effective in a high dimensional space and is compatible with different kernel functions specified for the decision function, with common kernels well provided and the freedom to specify custom kernels. More details about the theory of SVM can be referred in [16–18].

Most of the problems encountered in reality require the discrimination for more than two categories, and how to effectively extend SVM for multiclass classification is still an ongoing research issue. To solve these problems, a number of classification models were proposed, such as one-versus-rest approach [19], one-versus-one approach [20], direct multiclass SVM [21,22], and decision directed acyclic graph SVM (DAGSVM) [23]. In addition to these approaches above, some variants such as reordering DAGSVM [24], binary decision tree SVM [25], and error correcting codes SVM [26] have appeared in the literature. These novel methods have a wide range of applications including speech recognition, fault diagnosis, system discrimination, financial engineering and bioinformatics, etc. Table 1 shows the common multiclass SVM methods, and their computational complexity and advantages and disadvantages are compared [27]. In this paper, we use an Adaptive DAG (ADAG) [28]. This method improves the accuracy and reliability of DAGSVM and avoids the error accumulation, especially in the case of data sets with a large number of classes. A more detailed description of DAGSVM and ADAG is summarized as follows.

**Table 1.** A summary of the common multiclass SVM methods.

| Various Methods | Training Time | SVs | Size | Advantages | Disadvantages |
|---|---|---|---|---|---|
| One-versus-rest | Short | $K$ | Large | Simple Effective | Misclassification Rejected classification |
| One-versus-one | Long | $K*(k-1)/2$ | Moderate | High accuracy Fast classification | Inseparable problem Tend to overfit |
| Direct MSVM | Long | No SVs | Large | Natural optimization | Complex computation |
| DAGSVM | Long | $K*(k-1)/2$ | Large | Efficient to train and evaluate Avoid misclassification No rejected classification | Error accumulation |

### 2.2. Decision DAGSVM

The most primary problem with both one-versus-rest and one-versus-one support vector machine is unclassifiable regions. The binary-class linear classification model is defined as:

$$Y_{ij}(x) = w_{ij}^T \varphi(x) + b, \tag{1}$$

where $w_{ij}$ is a $y$-dimensional vector. $\varphi(x)$ is a mapping function that maps $x$ into the $y$-dimensional feature space. $b_{ij}$ is the bias tern, and $Y_{ij}(x) = -Y_{ji}(x)$. The unclassifiable region is shown in Figure 1 with three-classes, and its pairwise formulation is given by

$$R_i = \{x \mid Y_{ij}(x) > 0, j = 1, 2, ..., n, j \neq i\}. \tag{2}$$

$x$ will be classified into class $i$ if $x$ is in $R_i$. Otherwise, $x$ is classified by voting into the class with the most votes. Fot the input vector $x$, sign function is used to get the appropriate symbol. For real $x$,

$$sign(x) = \begin{cases} 1, & x > 0, \\ 0, & x = 0, \\ -1, & x < 0. \end{cases} \tag{3}$$

Here, we define $sign(x) = 1$ when $x = 0$. $Y_i(x)$ will be calculated accorading to the $sign(x)$ as follows:

$$Y_i(x) = \sum_{j=1,j\neq i}^{n} sign(Y_{ij}(x)), \tag{4}$$

and then $x$ will be classified into the appropriate class via

$$\underset{i=1,2,...n}{arg\ max}\ Y_i(x). \tag{5}$$

If $x \in R_i, k \neq i, Y_i(x) = n-1$ and $Y_k(x) < n-1$. In this case, $x$ is classified into $i$. However, if all $Y_i(x) \neq n-1$, plural *is* may appear in Equation (5), and $x$ is unclassifiable. $Y_i(x) = 0\ (i = 1,2,3)$ in the brown area in Figure 1, so this region is unclassifiable.
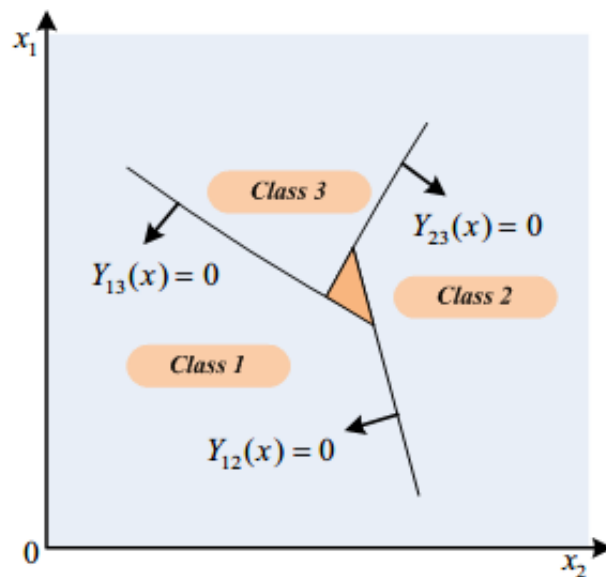


**Figure 1.** Unclassifiable regions.

Platt et al. proposed a decision DAG, which combined many two-class classifiers into a multiclass classifier [23]. It is a graph whose edges have an orientation and no cycles. For a $k$-class problem, DAG has $k$ leaves labeled by the classes, where each of the $k * (k-1)/2$ nodes is labeled with an element of a Boolean function. All nodes are arranged in a triangular shape with the single root node at the top, two nodes in the second layer and so on until the final layer of $k$ leaves. DAG starts at the root node, and the binary function at a node is evaluated. According to the value of binary function, the node is classified into the left edge or right edge. DAG is done until the list contains only one class. The current state of the list is the total state of the system. After the implementation of DAG, the unclassifiable region is resolved as shown in Figure 2.
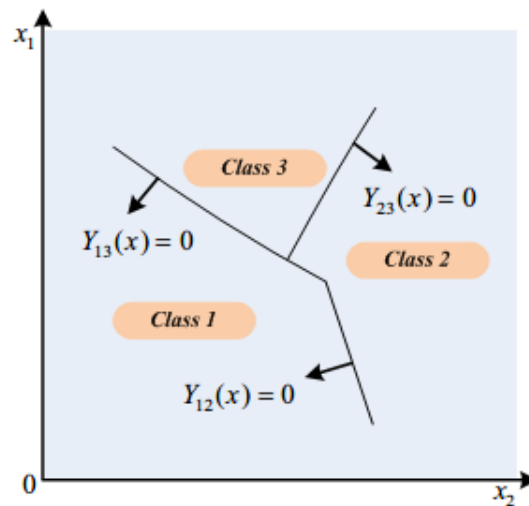
**Figure 2.** Directed acyclic graph classification.

### 2.3. ADAG

The decision DAGSVM can resolve the unclassifiable problem well. However, error accumulation is the main problem, and another drawback is that the final output is highly dependent on the ordering of the nodes. In this case, an adaptive DAG method is proposed to reduce the dependency on the sequence of nodes and lowers the depth of the DAG, and, consequently, the number of node evaluations for a correct class.

ADAG uses a decision tree with a reversed triangular structure in the testing stage. Training of a ADAG is the same as DAGSVM. In a $k$-class problem, it still contains $k * (k - 1) / 2$ binary classifiers and has $k - 1$ internal nodes. An example is shown in Figure 3, and the nodes are arranged in a reversed triangle with $k/2$ nodes at the top, $k/2^2$ nodes in the second layer and so on until the lowest layer of a final node.
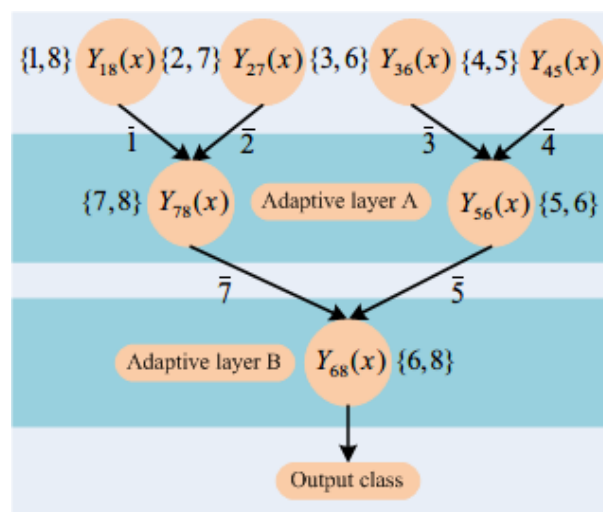


**Figure 3.** Adaptive directed acyclic graph classification.

ADAG will start at the top level, and the binary function at each node is evaluated. Like the DAGSVM, the node is then exited via the outgoing edge with a message of the preferred class. With the continued operation, the number of candidate classes will be reduced by half in each round, and then the ADAG process repeats until reaching the final node at the lowest level.

*2.4. Model Selection*

Selection of the appropriate kernel function is the main challenge in real applications. In this paper, we use the RBF kernel for the SVM classifier. The following are the main advantages of the RBF kernel:

- It can handle the case when the relation between class labels and attributes is nonlinear.
- It has fewer hyperparameters, which influences the complexity of model selection.
- It has fewer numerical difficulties.

$C$ and $\gamma$ are the two key parameters for an RBF kernel. However, it is not easy to find the best parameters. Through the practice of a large number of researchers, the method of finding the best two parameters was still to try exponentially growing sequences in a certain area (for example, $C = 2^{-5}, 2^{-4}, ..., 2^5$ and $\gamma = 2^{-5}, 2^{-4}, ..., 2^5$ ). A coarse grid is used to find a "better" region first, and then the process repeats until obtaining the best $(C, \gamma)$. The main drawback of this method is that different $C$ and $\gamma$ are likely to correspond to the highest accuracy. The selected parameters can improve the accuracy of validation data, but may cause over-learning state, and the final test set accuracy may not be very satisfactory because of the high penalty parameter. In this paper, we propose an improved grid-search approach, as shown in Figure 4. For example, the best $(C, \gamma)$ of motion 2 is $(1, 0.0313)$. $C$ and $\gamma$ with the smallest penalty parameter will be the best selection, and then the whole training set is trained again to generate the final classifier. Moreover, five-fold cross-validation, where we trained models on the first five parts of the whole dataset were selected randomly and tested with others. After five times of cross-validation, the average accuracy is obtained to avoid the overfitting problem as a result.
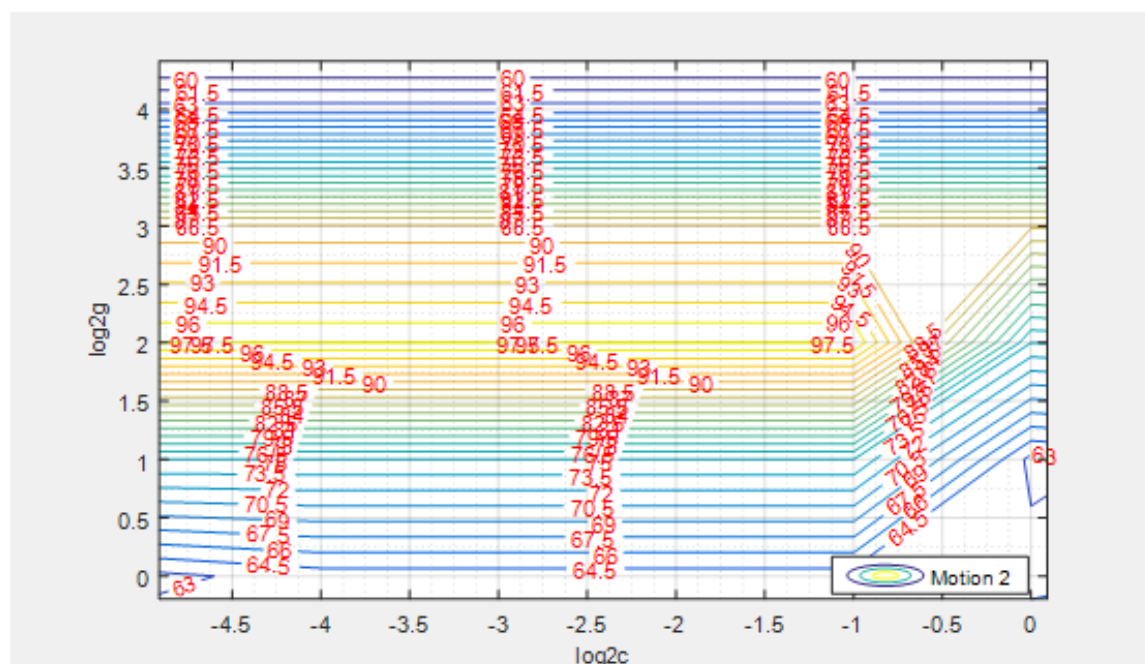


**Figure 4.** Improved grid search.

## 3. Hand Motion Capture System

The multiple sensor based hand motion capture system consists of a CyberGlove (Meta Motion, San Francisco, CA, USA) , a wireless tactile force measurement system from FingerTPS (PPS, Los Angeles, CA, USA) and a high frequency EMG capture system with Trigno Wireless Sensors (Delsys, Natick, MA, USA). It can capture the finger trajectories, the contact force signals, and the muscle signals

simultaneously. Next, the system architecture and the preprocessing module, including the hardware based synchronization and segmentation, will be presented, followed by the data capturing in the end.

### 3.1. System Configuration and Synchronization

The EMG capture system uses the Trigno wireless sensors shown in Figure 5a, and has 16 EMG channels and 48 accelerometer analog channels for motion capture. Each EMG sensor has a built-in trixial accelerometer, a guaranteed transmission range of 40 m and a rechargeable battery lasting a minimum of 7 h. It employs four silver bar contacts for detecting the EMG signal at the skin surface. The top of the sensor is shaped with an arrow to aid in the determination of this orientation. The arrow should be placed parallel to the muscle fibers underneath the sensor. In order to reduce the effects of crosstalk and redundancy, the sensor should also be placed in the center of the muscle belly away from tendons and the edge of the muscle. The signal resolution is 16 bits and a sampling rate of up to 4000 Hz.

CyberGlove I shown in Figure 5b is a fully instrumented glove, which provides up to 22 high-accuracy joint-angle measurements. The CyberGlove system has three flexion sensors per finger, four abduction sensors, a palm-arch sensor, and sensors to measure flexion and abduction. Each sensor is extremely thin and flexible, being virtually undetectable in the lightweight elastic glove. It uses proprietary resistive bend-sensing technology to accurately transform hand and finger motions into real-time digital joint-angle data. VirtualHand® Studio software (CyberGlove Systems, San Jose, CA, USA) is used to convert the data into a graphical hand, which mirrors the subtle movements of the physical hand. The sensor resolution is 0.5 degrees, the repeatability is one degree, and the sampling rate is 150 Hz.

The FingerTPS system utilizes highly sensitive capacitive-based pressure sensors to reliably quantify forces applied by the human hand, as shown in Figure 5c. It is the only practical and comfortable sensor solution that also connects wirelessly to your PC. By using powerful new Chameleon TVR Software (version 2012, PPS, Los Angeles, CA, USA), FingerTPS systems can be easily reconfigured and recalibrated for different uses of the hands on the fly. The data rate of the sensors is 40 Hz. The sensors have a scale range of 10–50 lbs and a sensitivity of 0.01 lbs. Video images can be captured and displayed in real-time, synchronized with tactile data.
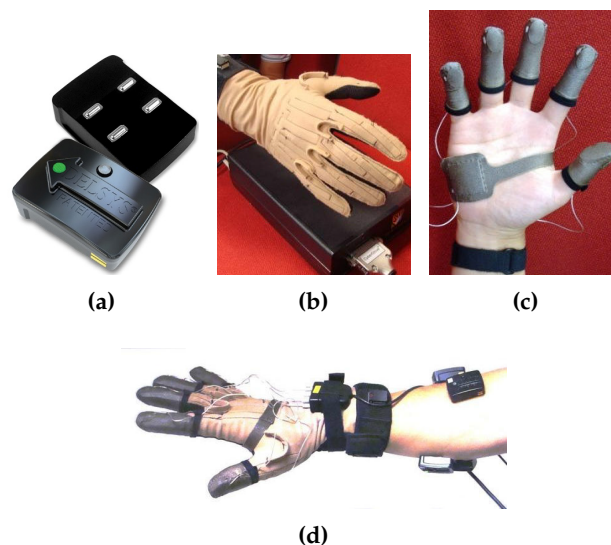


(a)　　　　　　　　(b)　　　　　　　　(c)



(d)

**Figure 5.** (**a**) Trigno wireless electromyography sensor; (**b**) FingerTPS; (**c**) CyberGlove I; (**d**) Multi-sensors for one hand.

A high-speed digital signal processor (DSP) has been used to acquire, process and send raw synchronized information digitally to a PC for analysis. The CPU speed of the DSP is far greater than 10 MHz for fast and efficient data acquisition. We utilize USB whose maximum data transfer rate is 10 megabits per second to realize the interface connection between the DSP and the computer. The resolution is set to 16 bits, and the three devices are sampled simultaneously.

### 3.2. Motion Segmentation

Motion segmentation is the key issue for separating the current motion with the next motion in the same type. Hendrich et al. realized the segmentation of the whole manipulation traces into several phases corresponding to individual basic patterns based on the finger forces [29]. By extracting the features of the multimodal data collected from human demonstrations of manipulation tasks, segmentation of the action phases and trajectory classification was accomplished [30]. Hand manipulation tasks typically take the intermediate state with short periods of time as a flag for the start and end of the action. Hence, we define the flat hand with no strength as the intermediate state. The motion begins when the finger angles change from the intermediate state, and ends when the finger angles change to the intermediate state. In this way, we utilize five-quick-grasp generated in the experiments to segment the motions when one type of the motions is finished and the participants are performing the next type of motions. It will enable muscles to contract five times, and enable glove and TPS to record up to five maximum in the trajectories simultaneously. However, even with training, participants may have difficulty fulfilling the tasks in one go during the experiments. There can be a situation in certain environments that may make some motions fail. For instance, the object may slip when grasping or being manipulated in-hand by accident. We design a protocol that, if the motion fails, the participants need to do a four-quick-grasp that indicates the motion captured before is invalid. The motion before the four-quick-grasp will be deleted during the separation process.

### 3.3. Motion Capturing

Eight healthy right-handed participants including two women and six men were invited to participate the experiments. All participants gave informed consent prior to the experiments and the ethical approval for the study was obtained from the University of Portsmouth Creative and Cultural Industries (CCI) Faculty Ethics Committee. Ten commonly hand motions were selected for experimental data sampling (shown in Figure 6). All participants were trained to manipulate different objects. The right way to grasp or manipulate the ten experimental motions were demonstrated before the participants performed. The contact points had been decided for every object, and every motion lasted about 2 to 4 s. Each motion was repeated 10 times. Between every two repetitions, participants had to relax the hand for 2 s in the intermediate state, which was opening the hand naturally without any muscle contraction. These intermediate states were used to segment the motions. To overcome the effects of muscle fatigue, once one motion with ten repetitions was finished, participants had to relax the hand for 2 min before the next motion started.
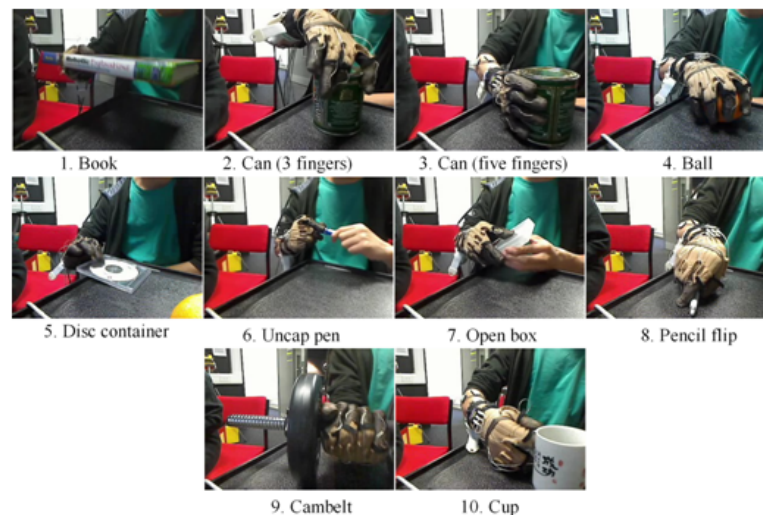
**Figure 6.** Hand motions including grasps and in-hand manipulations.

## 4. Experiment and Validation

We evaluate the proposed method by the experimental device including ten types of hand motions as mentioned, in which the whole raw data is divided into ten equal parts, part of which are used for training the models/templates and the rest for testing the algorithms. These recognition rates of ADAG will be compared from three aspects, including different sensor bases, different multiclass SVMs and other traditional methods, in order to fully confirm its performance.

### 4.1. Different Sensor Based Recognition

In our research, we assess the recognition rates of ten hand motions captured from the combined sensors, as shown in Figure 7. The proposed method gives a high average recognition rate of 94.57%, indicating the capability of the proposed hand motion recognition algorithm. The recognition rates are different for each motion. Specifically, we find that the ADAG method presented a perfect performance when identifying motion 1 (grasp and lift a book using five fingers with the thumb abduction) and motion 3 (grasp and lift a can full of rice using five fingers with the thumb abduction). For motion 2 (grasp and lift a can full of rice using thumb, index finger, and middle finger only), motion 4 (grasp and lift a big ball using five fingers), motion 5 (grasp and lift a disc container using thumb and index finger only), motion 6 (uncap and cap a pen using thumb, index finger, and middle finger), motion 7 (open and close a pen box using five fingers) and motion 8 (pick up a pencil using five fingers, flip it and place it on the table), all of the accuracies are up to 90%. In this case, the algorithm also reveals an excellent performance.

There are some reasons for good results. Firstly, more feature information is extracted to better capture hand movements for recognition. Secondly, five-quick-grasp and four-quick-grasp models identified using peak-detection algorithm, are utilized in the experiments to protect the integrity of raw data. Thirdly, the optimal kernels and parameters obtained will be conducive to gaining a better recognition rate, as well as the standard deviation of the accuracy.

**Figure 7.** Confusion matrix for the ten motions using ADAG, where the total accuracy is 94.57%.

Figure 8 shows the comparative experimental results and their variances of eight subjects using different sensor types. For different subjects performing all motions, the EMG based has the lowest average recognition rate of 85.55%, the data glove based average recognition rate is 88.30%, while the fingerTPS based average recognition rate is 89.20%. Compared with the results by using a single sensor, multi-sensor based has the highest accuracy. However, the recognition time increased because the extracted keypoints increased from hand motions. On the other hand, the identification of motion 9 (hold and lift a dumbbell) and 10 (grasp and lift a cup using thumb, index finger, and middle finger) also performs well, but it indicates the relatively lower discrimination compared with other motions. The main reason may be caused by the noise when collecting the EMG and finger angle information during two-hand cooperation and complex in-hand manipulation.
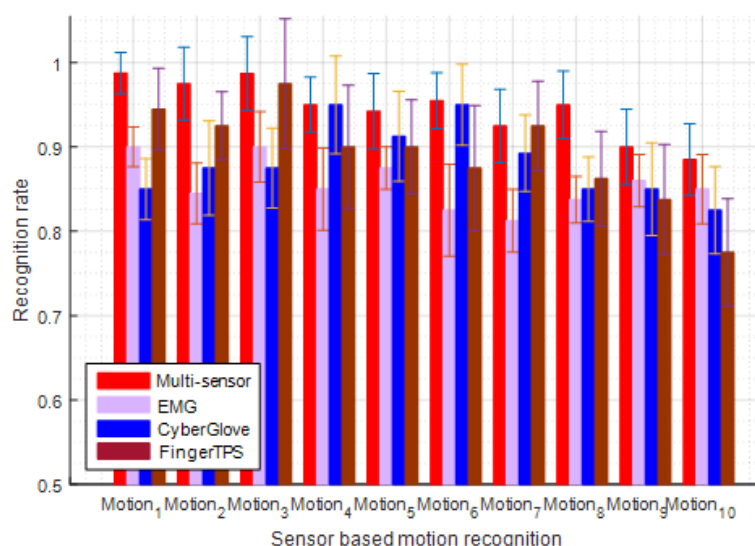


**Figure 8.** Comparative experimental results.

The recognition rates of different motions, as well as different sensor types, have been analysed through the above experiments. Different subjects exhibit significant individual variation in the same hand motion, such as the diversities of the applied acceleration, force, etc. Hence, it is necessary to

analyse the effects on the motion acquisition and recognition. Figure 9 describes the identification results of the same motion based on different subjects from the combined sensors. For different subjects performing all motions, we can clearly see that the difference of recognition rate is very apparent. The majority of them can get the average recognition rate of up to 90%, mainly caused by less training sample size and correct manipulation following the model. However, hand motion recognition from subject 7 has the worse recognition rates due to the misclassification of motions 5 and 6 with 22% and 46% error rates, respectively. The main reason for this result is the non-standard manipulation of motions 5 and 6. Hence, the influence of different subjects on the final experimental results can not be ignored.
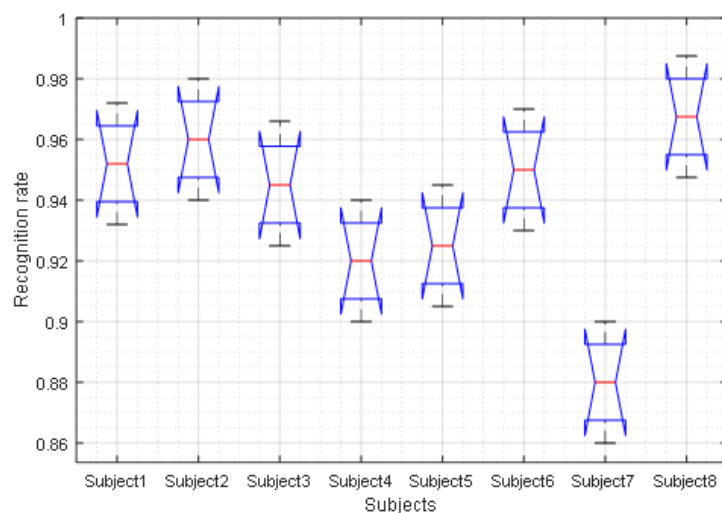


**Figure 9.** Recogniton rate with means of different subjects.

## 4.2. Different Multiclass SVM Based Recognition

In order to further verify the actual result of ADAG, it is very necessary to compare with other classical methods of multiclass SVM. We choose three other methods to analyze the same experimental data, and the corresponding results are shown in Figure 10. Due to different motions having different degrees of difficulty in manipulating, the recognition rate is gradually in decline with a slight fluctuation. The local similarity of matching features within a certain time frame increases the difficulty of motion recognition. Through the experimental comparison of four methods, it is increasingly clear that the one-versus-rest method has the lowest average recognition rate of 91.22%, the average recognition rate of ADG is 92.48%, while the one-versus-one has a higher average recognition rate of 94.51%, which is close to the average recognition rate of ADAG. Actually, the basic theories of the designs based on one-versus-one and ADAG are close to each other. They both evaluate all possible pairwise classifiers and thus induce $k * (k - 1) /2$ individual binary classifiers. The differences between them rely on the fact that the one-versus-one method uses the "maximum voting" to determine the class of unknown samples, while ADAG forms a tree-like structure to facilitate the testing phase. The comparison of different methods shows obvious advantages of the improved method and its remarkable significance, and the recognition accuracy becomes clearer with the increase of the number of classes.
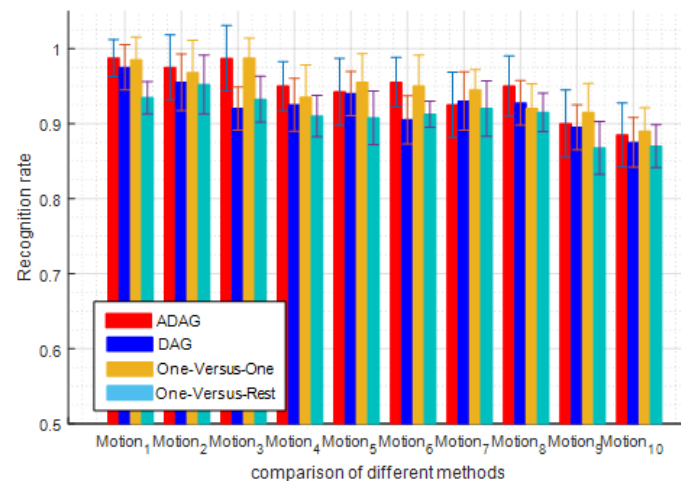
**Figure 10.** Recognition rate of different multiclass support vector machine methods.

## 4.3. Comparison of Classical Approaches

Neural networks (NN) are information processing systems for time-varying data analysis, and have been developed as generalizations of mathematical models of human cognition or neural biology based on the assumptions. The network consists of three layers, input layer with 16 nodes, hidden layer with 150 nodes, and output layer with 10 nodes, which corresponds to 10 recognized motions. Identity function for the NN activation functions has been employed to convert the net input to an output unit that is a binary signal. Fuzzy c-means (FCM) clustering is an unsupervised technique that has been successfully applied to feature analysis, clustering and classifier designs. In this paper, the set of feature vectors are clustered for subsequent use in the hand motion capture system. Once the clusters have been created, they are labeled manually. To evaluate the performance of ADAG for classifying different motions, ADAG is compared with both classical NN and FCM. Figure 11 presents the recognition rates of different methods for the different motions. FCM has the lowest average recognition rate of 78.94%. It turns out that FCM has a high fluctuation from other methods. The average recognition rate of NN is 89.20%, for which motion 4 and motion 7 have a higher accuracy than their neighbors. In terms of different methods and subjects, ADAG always presents the best performance.
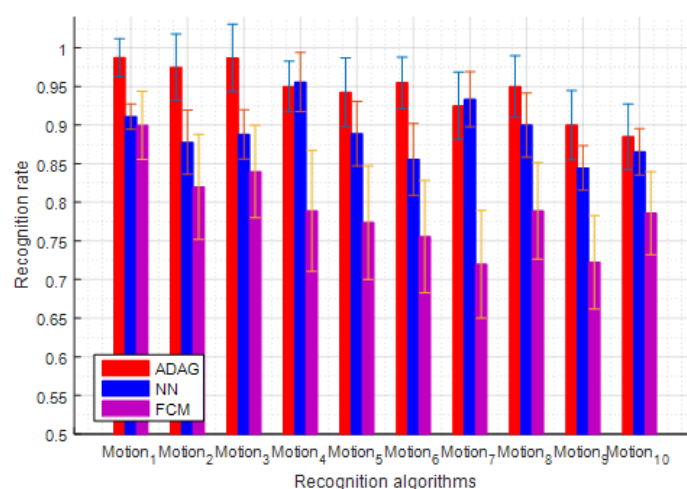


**Figure 11.** The comparison of three methods.

## 5. Conclusions

As mentioned above, we presented a multiple sensor based hand motion capture system for hand motion recognition. By applying proposed multimodal sensing technology, the EMG signals, the finger angle trajectories, and the contact force are simultaneously captured at a fast sampling rate. We perform manipulation primitives' segmentation of different motions using five-quick-grasp and four-quick-grasp models, which finally guarantee the validity of the raw data and reduce the number of repeats caused by failed action. In the hand motion recognition module, we firstly describe the current multiclass SVM methods and analyze their advantages and disadvantages, and then ADAG has been applied to recognize these ten motions from eight different subjects. From the experimental data, the recognition rate with multiple data fusion is up to 94.57%, which is much higher than individual sensor based. The feasibility and validity of this method also outperformed NN, FCM and three other traditional multiclass methods. Hence, multiple sensor based data fusion can merge rich information simultaneously to provide more accurate perception and make an optimal decision. In future work, we will implement robust and real-time hand motion recognition captured from vision based and contact based capturing systems, and further integrate them with robots to serve as a friendly and natural human–machine interaction and prosthetic hand control.

**Author Contributions:** Y.X. mainly conducted the background research and approach design, as well as the result validation and analysis. Z.J. finished the data acquisition and improved the writing. K.X., J.C. and H.L. provided important suggestion about the content and organisation of this paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Liu, J.; Luo, Y.; Ju, Z. An interactive astronaut-robot system with gesture control. *Comput. Intell. Neurosci.* **2016**, *2016*, 7845102, doi:10.1155/2016/7845102.
2. Saudabayev, A.; Varol, H.A. Sensors for Robotic Hands: A Survey of State of the Art. *IEEE Access* **2015**, *3*, 1765–1782.
3. Hichert, M. User Capacities and Operation Forces: Requirements for Body-Powered Upper-Limb Prostheses. Ph.D. Thesis, Delft University of Technology, Delft, The Netherlands, 24 February 2017.
4. Ju, Z.; Liu, H.; Xiong, Y. Fuzzy Empirical Copula for Estimating Data Dependence Structure. *Fuzzy Syst.* **2014**, *16*, 160–172.
5. Luzanin, O.; Plancak, M. Hand gesture recognition using low-budget data glove and cluster-trained probabilistic neural network. *Assem. Autom.* **2014**, *34*, 94–105.
6. Li, S.; Chu, Z.; Wang, Q.; Meng, F.; Wang, Y.; Ge, Y. Modeling of a 3D force flexible tactile sensor and its structural optimization. In Proceedings of the 2016 IEEE 11th Conference on Industrial Electronics and Applications, Hefei, China, 5–7 June 2016; pp. 2158–2297.
7. Al-Timemy, A.H.; Bugmann, G.; Escudero, J.; Outram, N. Classification of finger movements for the dexterous hand prosthesis control with surface electromyography. *J. Biomed. Health Inform.* **2013**, *17*, 608–618.
8. Metcalf, C.D.; Notley, S.V.; Chappell, P.H.; Burridge, J.H.; Yule, V.T. Validation and application of a computational model for wrist and hand movements using surface markers. *Biomed. Eng.* **2008**, *55*, 1199–1210.
9. Han, J.; Shao, L.; Xu, D.; Shotton, J. Enhanced computer vision with microsoft kinect sensor: A review. *Cybernetics* **2013**, *43*, 1318–1334.
10. Ju, Z.; Ji, X.; Li, J.; Liu, H. An Integrative Framework of Human Hand Gesture Segmentation for Human–Robot Interaction. *IEEE Sens. J.* **2015**, *99*, 1–11, doi:10.1109/JSYST.2015.2468231.
11. Park, J.; Lee, Y.; Hong, J.; Ha, M.; Jung, Y.-D.; Lim, H.; Kim, S.Y.; Ko, H. Giant tunneling piezoresistance of composite elastomers with interlocked microdome arrays for ultrasensitive and multimodal electronic skins. *ACS Nano* **2014**, *8*, 4689–4697.

12. Wong, R.D.P.; Hellman, R.B.; Santos, V.J. Haptic exploration of fingertip-sized geometric features using a multimodal tactile sensor. *Int. Soc. Opt. Photonics* **2014**, doi:10.1117/12.2058238.

13. Ju, Z.; Liu, H. Human hand motion analysis with multisensory information. *IEEE Trans. Mechatron.* **2014**, *19*, 456–466.

14. Marin, G.; Dominio, F.; Zanuttigh, P. Hand gesture recognition with leap motion and kinect devices. In Proceedings of the 2014 IEEE International Conference on Image Processing , Paris, France, 27–30 October 2014; pp. 1565–1569.

15. Fang, Y.; Hettiarachchi, N.; Zhou, D.; Liu, H. Multi-modal sensing techniques for interfacing hand prostheses: A review. *Sensors* **2015**, *15*, 6065–6076.

16. Burges, C.J.C. A tutorial on support vector machines for pattern recognition. *Data Min. Knowl. Discov.* **1998**, *2*, 121–167.

17. Abe, S. *Support Vector Machines for Pattern Classification*; Springer: London, UK, 2005.

18. Ben-Hur, A.; Weston, J. A user's guide to support vector machines. *Data Min. Tech. Life Sci.* **2010**, *609*, 223–239.

19. Vapnik, V. *The Nature of Statistical Learning Theory*; Springer Science & Business Media: New York, NY, USA, 2013.

20. Knerr, S.; Personnaz, L.; Dreyfus, G. Single-layer learning revisited: A stepwise procedure for building and training a neural network. In *Neurocomputing*; Springer: Berlin/Heidelberg, Germany, 1990; pp. 41–50.

21. Weston, J.; Watkins, C. Support vector machines for multi-class pattern recognition. *ESANN* **1999**, *99*, 219–224.

22. Crammer, K.; Singer, Y. On the algorithmic implementation of multiclass kernel-based vector machines. *Mach. Learn. Res.* **2001**, *2*, 265–292.

23. Platt, J.C.; Cristianini, N.; Shawe-Taylor, J. Large margin DAGs for multiclass classification. In Proceedings of the 12th International Conference on Neural Information Processing Systems, Denver, CO, USA, 29 November–4 December 1999; MIT Press: Cambridge, MA, USA, 1999; pp. 547–553.

24. Phetkaew, T.; Kijsirikul, B.; Rivepiboon, W. Reordering adaptive directed acyclic graphs: An improved algorithm for multiclass support vector machines. In Proceedings of the International Joint Conference on Neural Networks, Portland, OR, USA, 20–24 July 2003; Volume 2, pp. 1605–1610.

25. Madzarov, G.; Gjorgjevikj, D.; Chorbev, I. A Multi-class SVM Classifier Utilizing Binary Decision Tree. *Informatica (Slovenia)* **2009**, *33*, 225–233.

26. Passerini, A.; Pontil, M.; Frasconi, P. New results on error correcting output codes of kernel machines. *Neural Netw.* **2004**, *15*, 45–54.

27. Hsu, C.-J.; Lin, C.-W. A comparison of methods for multiclass support vector machines. *IEEE Trans. Neural Netw.* **2002**, *13*, 415–425.

28. Kijsirikul, B.; Ussivakul, N. Multiclass support vector machines using adaptive directed acyclic graph. In Proceedings of the 2002 International Joint Conference on Neural Networks, Honolulu, HI, USA, 12–17 May 2002; Volume 1, pp. 980–985.

29. Hendrich, N.; Klimentjew, D.; Zhang, J. Multi-sensor based segmentation of human manipulation tasks. In Proceedings of the IEEE Conference on Multisensor Fusion and Integration for Intelligent Systems, Salt Lake City, UT, USA, 5–7 September 2010; pp. 223–229.

30. Ju, Z.; Gao, D.; Cao, J.; Liu, H. A novel approach to extract hand gesture feature in depth images. *Multimedia Tools Appl.* **2016**, *75*, 11929–11943.