A Leap Password based Verification System

Aman Chahar *, Shivangi Yadav *, Ishan Nigam, Richa Singh, Mayank Vatsa
IIIT-Delhi, New Delhi, India
{aman11014, shivangi11104, ishann, mayank, rsingh}@iiitd.ac.in

Abstract

Recent developments in three-dimensional sensing devices has led to the proposal of a number of biometric modalities for non-critical scenarios. Leap Motion device has received attention from Vision and Biometrics community due to its high precision tracking. In this research, we propose Leap Password; a novel approach for biometric authentication. The Leap Password consists of a string of successive gestures performed by the user during which physiological as well as behavioral information is captured. The Conditional Mutual Information Maximization algorithm selects the optimal feature set from the extracted information. Match-score fusion is performed to reconcile information from multiple classifiers. Experiments are performed on the Leap Password Dataset, which consists of over 1700 samples obtained from 150 subjects. An accuracy of over 81% is achieved, which shows the effectiveness of the proposed approach.

1. Introduction

The emergence of specialized three-dimensional sensors to track humans has opened up avenues in biometric research to implement novel recognition systems. Such systems are ideally suited towards non-critical identification applications such as accessing public workplaces, amusement parks, and entering buildings. Development of sensors such as Microsoft Kinect and Leap Motion has aided the proposal of new authentication approaches as well as augmentation of prior approaches. For example, face recognition has now been studied extensively using Microsoft Kinect [4, 7, 10, 13].

A Leap Motion device tracks a user's hand in three-dimensional space at a spatial resolution of 10^{-4} m and a temporal resolution of 120 frames per second using infrared sensors. The device can be used to capture the spatiotemporal trajectories of both hands of a user and thus, has applications in gesture recognition, handwriting analysis, and biometric verification. Vikram et al. [11] track the moving patterns of fingers and apply this information to handwriting recognition. Recently, Nigam et al. [9] have proposed a biometric verification system, termed as the Leap signature. The 3D signature, described as a pattern rendered by the user in 3D space in absence of any tactile feedback or obstruction, is an exclusively behavioral biometric trait. Information from the 3D signature is captured by computing Histogram of Oriented Optical Flow and Histogram of Oriented Trajectory.

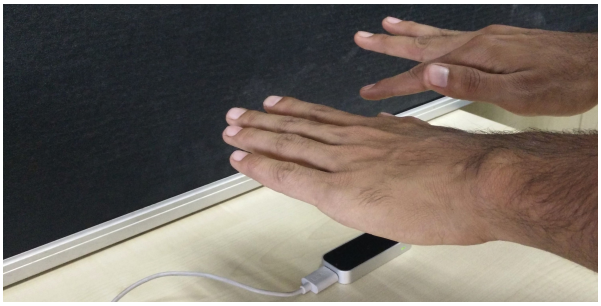

Figure 1: Acquisition of Leap password: the subject is performing fingertap gesture.

In this research, we explore the possibility of utilizing physiological as well as behavioral information captured by a Leap Motion device to verify the identity of the user. We propose that the combination of several uncorrelated descriptors such as time taken to perform successive characters of the gesture password (behavioral information), dimensions of the subject's fingers and palms (physiological information), and the password itself (memorized gestures) is likely to provide sufficient discriminatory information to be able to provide a high recognition performance.

In the proposed approach, the Leap Motion device is used to allow the user to enter a password string consisting of six gestures performed by using one of their ten fingers at a time. The extracted features are studied from the point of view of their relevancy and redundancy, and a subset of features is chosen which maximizes the conditional information. This information is classified independently and

* Equal contribution from A. Chahar and S. Yadav.

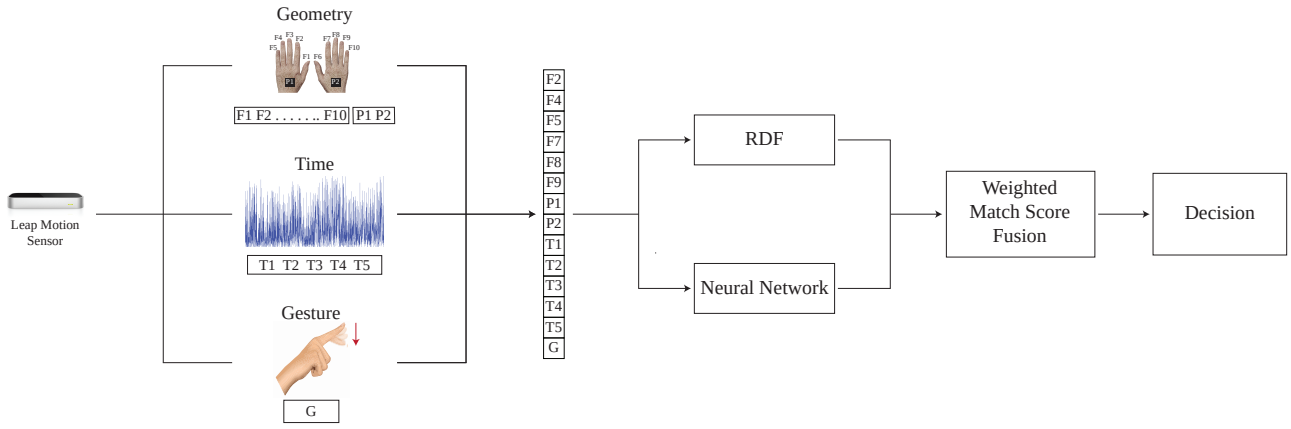

Figure 2: Summary of the proposed aLPhabet method for Leap Password verification.

match-score level fusion is performed. The key contribution of this research are:

1. *Leap Passwords*, 3D gesture based passwords which capture physiological and behavioral information using the Leap Motion device, are proposed as a new approach for person verification.
2. A novel algorithm, termed as aLPhabet, is proposed to encode information for the Leap Password.
3. Leap Password Dataset is collected which consists of Leap Passwords from 150 individuals. The dataset also contains samples of intentional spoofing attempts made by trained users.

2. Proposed aLPhabet Framework

The motivation for the proposed algorithm, termed as *aLPhabet*, is the observation that a combination of uncorrelated physiological and behavioral information can be highly discriminative and likely to provide a means to verify user's identity with sufficient confidence. Figure 2 illustrates the steps involved in aLPhabet algorithm.

2.1. Information Captured from Leap Motion Device

The Leap Motion device measures the physiological characteristics of the user's hands (Figure 2) at an extremely high spatial resolution and also registers the points in time when the successive gestures of the Leap Password are performed with high precision. The following information is captured by the Leap Motion device:

- Lengths of ten fingers and widths of two palms of the user, at a spatial resolution of 10^{-4}m.
- Time taken by the user in performing successive gestures of the Leap Password, at 120 frames per second.

2.2. Feature Extraction

From the raw features, three types of features are extracted and computed: (1) measurements of the fingers and palms of the user's hands, (2) time taken by the user in performing successive gestures in the Leap Password, and (3) a similarity measure for the Leap Password performed by the user compared to a gallery Leap Password. These features are discussed in detail below.

2.2.1 Dimensions of fingers and palms

The dimensions of the fingers and the palms (in millimeters) captured by the Leap Motion device are utilized as the first set of features. Since the precision of the Leap Motion device is very high and measurements are captured distinctly, any post-processing is not performed for the physiological information. The feature descriptor, $\vec{F}_{dim}$, is formed by concatenating the captured measurements:

$$\vec{F}_{dim} = [F_1, F_2, ..., F_9, F_{10}, P_1, P_2] \quad (1)$$

where, F_i refers to the length of the i^{th} finger and P_j refers to the width of the j^{th} palm.

2.2.2 Time taken between successive gestures

The time taken by a user to perform successive finger-tap gestures provides discriminatory behavioral information about the user. The Leap Password captures a string of 6 gestures performed by the user. The time taken by the user between successive gestures (in milliseconds) is recorded using the Leap Motion Software Development Kit, and the behavioral feature, $\vec{F}_{time}$, is formed by concatenating the captured measurements:

$$\vec{F}_{time} = [T_1, T_2, T_3, T_4, T_5] \quad (2)$$

where, T_i is the time taken by the user between gestures g_i and g_{i+1}.

2.2.3 Similarity measure for gesture password

The gesture password is a string of finger-tap gestures performed by the user through any of their 10 fingers, one finger at a time. This password is compared to the password of the claimed identity. The algorithm used to compute the similarity score is the popular string matching method, the *Levenshtein* algorithm [6].

Levenshtein Algorithm: Given two strings G_1 and G_2, the edit distance $G(G_1, G_2)$ is defined as the minimal series of edit operations that transform G_1 into G_2. The alphabet on which the strings are defined are 10 distinct characters - the fingers of the user. The following operations are permitted for transforming the strings:

1. *Insertion*: If $G_i = g_{i1}g_{i2}$, insertion of the symbol g_k produces the string $g_{i1}g_kg_{i2}$.
2. *Deletion*: If $G_i = g_{i1}g_kg_{i2}$, deletion of the symbol g_k produces the string $g_{i1}g_{i2}$.
3. *Substitution*: If $G_i = g_{i1}g_kg_{i2}$, substitution of the symbol g_k produces the string $g_{i1}g_lg_{i2}$, where $g_k \neq g_l$.

Due to the frequent mistakes made by the sensor in detecting taps and registering single taps as multiple taps, the penalty affixed for the *Substitution* operation is more than the *Insertion* and *Deletion* operations. The information extracted from comparing the gesture password strings is the scalar similarity measure, *Edit Distance (G)*:

$$F_{dist} = G \tag{3}$$

From equations 1 to 3, features extracted from Leap Password can be written as:

$$\vec{F}_{comp} = [\vec{F}_{dim},\ \vec{F}_{time},\ F_{dist}] \tag{4}$$

2.2.4 Entropy analysis of Leap Password

The Leap Password consists of 6 gestures. There are 10 possibilities (fingers) for performing every gesture in the Leap password. Hence, the total number of possible passwords is 10^6. The range of finger dimensions will be a function of the dataset's population; the range for the Leap Password Dataset is approximately 10mm, and the precision of the Leap Motion device is 10^{-4}mm. Hence, the number of possible finger dimensions is 10×10^4 for ten fingers. Similarly, the palm widths have a 20mm range; hence the number of possible palm dimensions is 20×10^4 for two palms.

The range of the time taken between gestures in the Leap Password Dataset is approximately 250 ms. The precision of the Leap Motion device is 10^{-2} ms. Five such timestamps are recorded during the gesture password. The total number of possible gaps in time is, thus, 1.25×10^5. Hence, the total number of possible Leap passwords is 5×10^{22}.

2.3. Feature Selection

The discriminative capability offered by a feature at the cost of increasing dimensionality is a critical trade-off. Feature selection improves prediction performance of classifiers, provides faster and more cost-effective predictors, and allows us to gain a better understanding of the underlying processes that govern the generation of data [5].

We study filter-based criteria to select features that are useful for classification. The advantage of using filter-based selection methods lies in their effective computation time and robustness to overfitting [2]. The selection criterion (J) is a measure of correlation between the feature (X_k) and the class label (Y); a strong correlation implies a greater predictive ability when the feature is used in the presence of the currently selected feature subset (S):

$$J(X_k) = I(X_k; Y|S) \tag{5}$$

In this research, we apply the Conditional Mutual Information Maximization criterion (J_{CMIM}), proposed by Fleuret [3]:

$$J_{CMIM}(X_k) = (X_k;Y) - max_{X_j \in S}[I(X_k;X_j) - I(X_k;X_j|Y)]$$

where, I represents *mutual information*, X_k represents a *feature*, Y represents *class label*, and S represents the set of features that have been selected prior to processing X_k.

From $\vec{F}_{comp}$, the feature vector computed from Leap Password, J_{CMIM} criterion selects relevant features and a new representation, $\vec{F}_{select}$, is selected based on the minimum misclassification accuracy obtained on the training set. In context of the proposed algorithm, we observe that the following features are selected:

$$\vec{F}_{select} = [F_2, F_4, F_5, F_7, F_8, F_9, P_1, P_2, \vec{F}_{time}, F_{dist}] \tag{6}$$

2.4. Classification and Fusion

The suitability of each class of features ($\vec{F}_{dim}$, $\vec{F}_{dist}$, $\vec{F}_{time}$) is explored independently as well as in combination with other classes of features ($\vec{F}_{comp}$) using the Naïve Bayes (NB), Neural Network (NNET), and Random Decision Forest (RDF) classifiers. Further, match-score level fusion is explored to analyze the effect of combining scores obtained from different classifiers such that,

$$Score_{fuse} = \sum_i W_i * score_i \tag{7}$$

where, $score_i$ is the score obtained from the i^{th} classifier and W_i is the associated weight.

3. Leap Password Dataset

Leap Motion based biometric authentication is relatively unexplored and to the best of our knowledge, the IIITD Leap Signature Dataset [9] is the only publicly available Leap Motion device dataset. However, as discussed previously, Leap Signatures do not capture any physiological subject information. Therefore, another major contribution of this research is the release of the Leap Password Dataset (LPD) [1]. This database consists of information captured from 150 subjects performing fixed-length gesture passwords on a Leap Motion device. The following information is recorded from samples provided by users:

1. Dimensions of the palms and fingers of user's hands.
2. Similarity measure for the password entered by the user and the user's password enrolled in the dataset.
3. Time taken to input successive gestures by the user.

In order to study the resilience of the proposed algorithm towards intentional impostor attempts, spoofing samples are also created by three different individuals: two males (one medium build, one large build) and one female (small build), their age ranging from 21 to 23 years. Each spoofing impostor is shown the complete gesture password for every subject and given sufficient time to acclimatize themselves with the password string and to perform the Leap Password. Each impostor records two spoofing attempts for every subject. Table 1 summarizes the characteristics of the dataset. The dataset consists of 2 subsets with no overlap between subjects: Leap Password Dataset Train (LPD Train) and Leap Password Dataset Test (LPD Test). Both subsets are captured in similar environments and under identical constraints. LPD Test consists of a higher number of samples per subject to extensively test the proposed system.

Table 1: The Leap Password Dataset (LPD)

Number of unique subjects	150
Samples per subject	5-12
Number of samples in dataset	1275
Number of spoofing attempts in dataset	450
Number of total samples in dataset	1725
Training Data (LPD Train)	
Number of subjects	75
Samples per subject	5
Number of total samples in dataset	375
Testing Data (LPD Test)	
Number of subjects	75
Samples per subject	12
Spoofing attempts per subject	6
Number of total samples in dataset	1350

4. Experimental Analysis

The performance of the proposed aLPhabet algorithm for biometric verification is analyzed on the Leap Password Dataset. As discussed in Section 2, the aLPhabet algorithm has the following stages: feature extraction, feature selection, classification, and fusion. Experiments are performed to analyze the relevance and redundancy contributed by each feature and an exhaustive analysis of the optimal recognition strategy is performed using Neural Network, Naïve Bayes, Random Decision Forest classifiers.

4.1. Analysis of Features and Classifiers

We begin by analyzing the features extracted from the Leap Motion device. Since the information descriptor consists of information from diverse sources including physiological and behavioral information, we consider classification using different algorithms, namely, Naïve Bayes (predictive classifier), neural network (discriminative classifier), and random decision forest (ensemble classifier). A comparative analysis of features taken independently as well as in combination is provided in Table 2. It is observed that geometry features yield better performance than other features. We also perform classification for pair-wise feature representations, but it is observed that these representations do not yield high accuracy. With all the features concatenated and RDF classifier 75.47% Genuine Accept Rate (GAR) is achieved at 1% False Accept Rate (FAR).

4.2. Analysis of Feature Selection

Since the basis of using feature selection algorithms is to maximize relevance and minimize redundancy, it is critical to assess the ranking of the Leap Password features provided by the CMIM criterion. It is:

$$T_1 > T_2 > T_3 > T_4 > ED > T_5 > F_7 > F_8 > P_1 > P_2 ...$$
$$... \; F_2 > F_4 > F_5 > F_9 > F_{10} > F_1 > F_3 > F_6 \qquad (8)$$

During exploratory analysis and during the preparation of the Leap Password Dataset, it is observed that the thumbs (F_6, F_{10}) are consistently mis-tracked while the index fingers (F_2, F_7) are consistently tracked by the Leap Motion device. CMIM ranks the two thumb features among the three lowest features extracted from the Leap Password. Since the physiological information for the fingers extracted from a user's left and right hand fingers is likely to be similar for corresponding pairs of fingers, it is observed that the five lowest features consist of atleast one finger corresponding to one out of the thumb, middle, ring, and little fingers.

Brown et al. [1] have proposed that the Joint Mutual Information (JMI) criterion [12] provides the best trade-

[1] http://www.iab-rubric.org/resources/lpd.html

Table 2: Genuine accept rates (%) at 1% false accept rates of independent feature types and their fusion.

Classifier	Individual Features			All features without CMIM	All features with CMIM	All features with JMI
	Time	Geometry	Gesture			
Naïve Bayes	43.39	67.29	29.13	71.29	**75.78**	66.75
Random Decision Forest	34.69	50.95	28.91	75.47	**78.04**	67.43
Neural Network	46.81	59.62	28.91	73.82	**78.55**	38.44

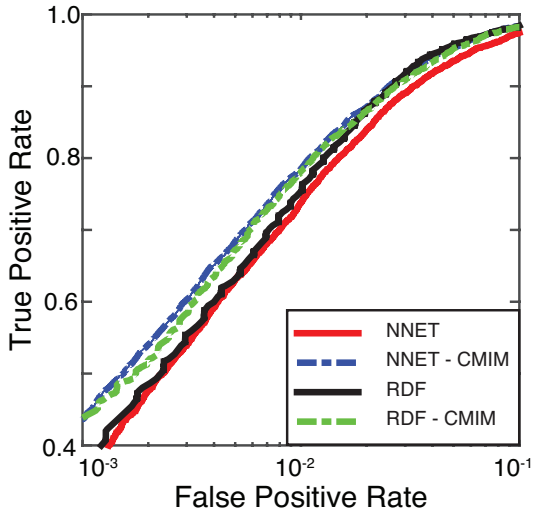

Figure 3: Performance of feature fusion and selection.

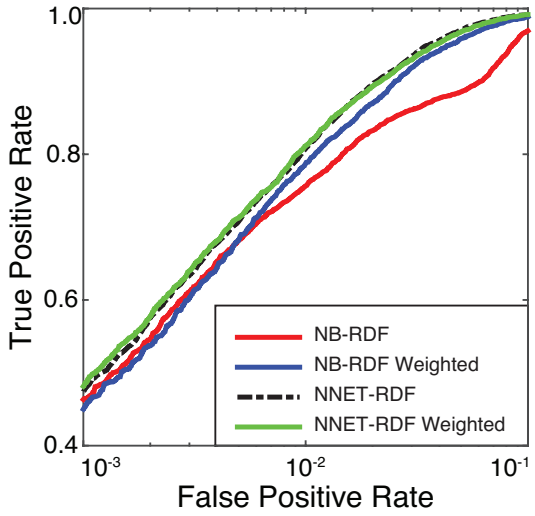

Figure 4: Performance of match-score fusion.

off in terms of accuracy and stability with small data samples. Experiments are performed to compare several feature selection criteria with JMI. It is observed that the Conditional Mutual Information Maximization (CMIM) criterion [3] outperforms JMI. A summary of the performance of CMIM and JMI algorithms is provided in Table 2 along with the recognition performance of the classifiers in the absence of feature selection.

4.3. Statistical Analysis

The McNemar test [8] provides a measure of the statistical correlation of the recognition performance of two classifiers. At 0.05 significance level, a $\chi^2 \geq 3.84$ suggests that the accuracies obtained by the two classifiers are statistically different. We analyze the above-mentioned classifiers in pairs of two classifiers each using the McNemar test (3) and compute the χ^2-scores as follows:

$$\chi^2 = \frac{(N_{01} - N_{10})^2}{N_{01} + N_{10}} \tag{9}$$

where, N_{01} and N_{10} represent the frequency of misclassifications made by one of the algorithms while the other algorithm classifies correctly. The χ^2-score values, as observed in Table 3, are above the threshold for the comparison of random decision forest with Naïve Bayes ($\chi^2 = 40.04$) and neural network classifiers ($\chi^2 = 34.17$). However, when comparing neural network with Naïve Bayes classifier, it fails to predict statistical independence.

4.4. Analysis of Score-level Fusion

From Table 2, it is observed that for the features selected by the CMIM criterion ($\vec{F}_{select}$) Naïve Bayes, random decision forest, as well as neural network classifiers yield similar recognition performance. As observed above, the combination of random decision forest with Naïve Bayes classifier and the combination of random decision forest with neural network classifiers is found to be statistically independent. Match-score fusion is performed for these classifiers in the following ways: sum score fusion, weighted sum score fusion, and SVM score fusion (scores are fed to a linear SVM and output is treated as a match score). Table 4 summarizes the results for these match-score fusion experiments. It is observed that a weighted sum of match-scores for random decision forest and neural network classifiers achieves a genuine accept rate of 81.17% at a false accept rate of 1% (Figure 5). The spoofing samples that are developed as part of the database are used to analyze the effect of attacks on Leap Password. We observe that 8% of spoofing

Table 3: Comparison of statistical correlation of classifiers - McNemar Test.

Decision	NNET +	NNET -
RDF +	396641	3160
RDF -	2712	2037
χ^2 score	34.17	

Decision	NNET +	NNET -
NB +	395385	4416
NB -	3841	908
χ^2 score	2.12	

Decision	NB +	NB -
RDF +	395496	3857
RDF -	3730	1467
χ^2 score	40.04	

attempts succeed at gaining access in the proposed system. The high number of successful spoofing attempts is likely caused due to (i) errors in finger detection in the Leap Motion device due to sensor level noise, and (ii) the assumption that the intruder has obtained the gesture password via *shoulder surfing*.

Table 4: GAR at 1% FAR for match-score level fusion.

	Sum Rule	Weighted Fusion	SVM Fusion
NB and RDF	75.64%	78.77	76.34
RDF and NNET	80.77	**81.17**	75.58

5. Conclusion and Future Work

This research introduces the Leap Password, a biometric authentication approach which combines physiological and behavioral information. The proposed algorithm, aLPhabet, utilizes the physiological and behavioral information captured via Leap Motion device. Features such as geometry of fingers and palms, time taken by the user to perform the Leap gesture password, and the gesture password are used for verification purposes. In conjunction, classifiers such as RDF and neural network are used for classification. On the prepared Leap Password Dataset, we achieve 81% GAR at 1% FAR. This shows the promise towards this novel biometric authentication approach. As a future research direction, we plan to study the possibility of incorporating additional information from the Leap Motion device to provide more discriminatory capabilities to the aLPhabet algorithm, and minimize spoofing attempts.

References

[1] G. Brown, A. Pocock, M.-J. Zhao, and M. Luján. Conditional likelihood maximisation: A unifying framework for information theoretic feature selection. *Journal of Machine Learning Research*, 13:27–66, 2012.

[2] W. Duch. *Feature Extraction: Foundations and Applications, Chapter 3*, volume 207. Springer-Verlag Berlin Heidelberg, 2006.

[3] F. Fleuret. Binary feature selection with conditional mutual information. *Journal of Machine Learning Research*, 5.

[4] G. Goswami, M. Vatsa, and R. Singh. RGB-D Face Recognition With Texture and Attribute Features. *IEEE Transactions in Information Forensics and Security*, 2014.

[5] I. Guyon and A. Elisseeff. An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3:1157–1182, 2003.

[6] V. I. Levenshtein. Binary codes capable of correcting deletions, insertions and reversals. *Soviet Physics Doklady*, 10(8):707–710, 1966.

[7] B. Li, A. Mian, W. Liu, and A. Krishna. Using Kinect for face recognition under varying poses, expressions, illumination and disguise. In *IEEE Winter Conference on Applications of Computer Vision*, pages 186–192, Jan 2013.

[8] Q. McNemar. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, 12(2):153–157, 1947.

[9] I. Nigam, M. Vatsa, and R. Singh. Leap signature recognition using HOOF and HOT features. In *21st IEEE International Conference on Image Processing*, 2014.

[10] M. Pamplona Segundo, S. Sarkar, D. Goldgof, L. Silva, and O. Bellon. Continuous 3D Face Authentication Using RGB-D Cameras. In *IEEE CVPR-Workshops*, pages 64–69, June 2013.

[11] S. Vikram, L. Li, and S. Russell. Writing and Sketching in the Air, Recognizing and Controlling on the Fly. In *Proceedings of CHI 2013 Extended Abstracts*, pages 1179–1184. ACM, 2013.

[12] H. H. Yang and J. E. Moody. Data Visualization and Feature Selection: New Algorithms for Nongaussian Data. In *Advances in Neural Information Processing Systems*, pages 687–702, 1999.

[13] Z. Zhang. Microsoft Kinect Sensor and Its Effect. *IEEE MultiMedia*, 2012.