Asian Journal of Convergence in Technology Volume III, Issue II
ISSN No.:2350-1146, I.F-5.11

Feature Extraction and Matching for different Intensity values using scale invariant feature transform.

Sandeep B. Patil
Research Scholar (PhD student)
FET, SSGI, SSTC, Bhilai.
patilsandeepb1212@gmail.com

Dr. R.H. Talwekar
Prof. Govt. college of Engineering,
Raipur
ursk.talwekar@gmail.com

Abstract

India, having less awareness towards the deaf and dumb peoples results in increase the communication gap between deaf and laborious hearing community. Sign language is often developed for deaf and laborious hearing peoples to convey their message by generating the various sign pattern. The Scale invariant feature Transform has been accustomed perform reliable matching between totally different image of the same object. This paper implements the assorted phases of scale invariant feature transform to extract the distinctive features from Devnagri sign language gestures. The intensity of the original image is varied and then the feature extraction and matching has been performed between original and intensity varied image. The experimental result shows the intensity of original image is changed by a pair of,3,4,and 5 times and the system achieves more than 99% of accuracy.

I. INTRODUCTION

In recent years, there are various analysis contributions within the field of Devnagri sign language, a part of Indian sign Language (ISL) that helps in identification of persons supported their traits or characteristics. This space isn't terribly straightforward as a result of no normal databases are available for the Devnagri Sign language. The importance of the matter is simply illustrated by victimization natural gestures applied beside verbal and nonverbal communication. The employment of hand gestures in support of verbal communication is extremely helpful for the folks having no visual contact. Completely different approaches for recognizing hand gestures are available in literatures; few of them need sporting marked gloves or attaching further hardware to the body of the subject. These approaches are less likely apply to the globe applications whereas vision-based approaches are consider as non-intrusive and thence a lot of doubtless to be used for globe applications.

Indian sign Language is an endeavor during this direction that deals with recognition of varied static and symbolic characters generated by hand gestures. The biometry uses the ISL gestures as human traits and acknowledges the characters consequently, that is extremely helpful for deaf and dumb peoples. Many professionals estimate that the deaf population in India is roughly 5 million and onerous hearing peoples is 10 millions. Around one hundred million peoples are associated and involved with those 15 million peoples caring them and serving to them in communication. The concerned peoples are family members, social employees, audiologist, skilled, academics etc. If any answer is developed for serving to the deaf and dumb peoples to acknowledge the language then it would be a big contribution towards society and humans.

II. LITERATURE REVIEW

A numerous research works on Sign language is being carried out and some of the researchers have already contributed in this area. Study of sign language was made and below is a few report of literature review:

Heera et al, 2017, made a sensor based glove system to converts hand gesture to speech. A Bluetooth module and an Android Smart Phone have been used to detect the hand gestures. The gloves have been used to produce artificial speech and environment similar to daily communication. The variety of sensors like flex, gyroscope and accelerometer sensors have been used to converts Indian Sign Language to speech. The said system also aims to integrating the results of the sensor with a smart phone, map the sensor reading to a corresponding sign and stored in a database. The system is autonomous, user friendly and a completely mobile system. The proposed system can be extended to the applications in the field of education under Virtual Reality. The gloves can also be used for interacting with set of electronic devices across house using centralized IoT hub [6]. Amitoj Singh et al, 2017, reported Punjabi Sign Recognition System to converts Indian sign language into text using LabVIEW software. The proposed system support for static gesture recognition where hand motion, orientation and mask of the image has created. The SIFT algorithm has been used to extract the stable features which further used to design a human computer interface system for recognizing sign language of the deaf and dumb accurately. The system is not applicable for real time images, as well as the time computation for various phases of SIFT has not been mentioned. The proposed system used for only two punjabi sign. The accuracy has not been reported [7]. Anup Kumar et al, 2016, suggested a system which can convert speech to sign language for American Sign Language. The algorithm is capable of extracting signs from video sequences under cluttered and dynamic background. Further the image is preprocessed using skin color segmentation. The said system is able to differentiate between static and dynamic gestures and classification has been done using Support Vector Machines. The standard module of speech recognition that is –Sphinx has been used. Experimental results show satisfactory segmentation of signs under diverse backgrounds and relatively. For 24 alphabet of ASL, 93% accuracy has been achieved. The system needs to be improved for unfavorable environment, needs to recognize more gestures like those involving two hands, also needs to deal with co-articulation [8]. Alhussain Akoum et.al, 2015, reported the hand gesture recognition approach for ASL Using hand extraction algorithm, the gestures has been recognized using various steps and is transformed to written character and speech. The SIFT algorithm has been used to extract features and matching has been done using 2-D Correlation Coefficient, Edge Detection and date histogram. The overall matching result is 85% - 90% has been mentioned for words and letters. The applied input gestures needs to be preprocessed, the same background is needed for better accuracy. The system has to process 1200 images for three independent algorithms and hence more execution time has been required for matching [9]. Nanivadekar et al, 2014, presented the Sign Language recogintion system. The database has been created by acquiring the videos from the signers moment. Hand tracking and Segmentation has been performed to extract features from a particular gesture. The results demonstrated working of motion tracking, edge detection and skin color detection individually and their combined effect. The said algorithm has advantage is that it includes dynamic gestures and phrases like "Thank you", "Sorry", "Help Me", "Danger" etc. It also included the phrases that involve facial movements in the gesture [10]. S. Pramada et.al, 2013, proposed Intelligent Sign Language Recognition Using Image Processing The hand segmentation technique uses background subtraction method to segment the hand with uniform and complex background from the image. The idea consisted of designing and building up an intelligent system using image processing, machine learning and artificial intelligence concepts to take visual inputs of sign language's hand gestures and generate easily recognizable form of outputs. D. J. Rios-Soria et.al, 2013, proposed another application of gesture language is human-computer interaction, which uses hand gestures as input data to a computer through webcam. In HCI, a visual interface is created to provide a natural way of communication between man and machine. The underlying algorithm utilizes only computer-vision techniques. The tool is able to recognize in real time six different hand gestures, captured using a webcam. classification, the execution time for recognition would be large. Ashwin S.Pol et. al, 2013, presented Sign Language Recognition Using Scale Invariant Feature Transform and SVM, The SIFT features were extracted for ASL and codes are assign to all features using K mean clustering. The SVM classifier has been used for classification. The proposed system used four postures of ASL viz. A, B, V and Five. The recognition accuracy for each postures are varies from 82% to 90% [11]. Goyal et. Al, 2013, developed Sign Language Recognition System for deaf and dumb People, the presented system used SIFT algorithm for feature extraction of ISL gestures. The mentioned 26 ISL gestures and 9 alphabets gives 95% of accuracy. The feature matching algorithm has not been mentioned [12]. Geetha M. et al, 2012, describes a novel vision-based recognition of Indian Sign Language Alphabets and Numerals using B-Spline. A boundary extracted algorithm has been used as a control point. Then the B-Spline curve is subjected to iterations for smoothening resulting in the extraction of Key Maximum Curvature points (KMCPs), which are the key contributors of the gesture shape. [13] Kishore P.V.V., et al, 2012, proposed an approach for a real time system to recognize gestures of sign language from videos under complex backgrounds. The proposed algorithm converts a video into a voice and text command. The system has been implemented successfully for 351 signs of Indian Sign Language under different possible video environments. The recognition rates are calculated for different video environments [14]. Jayashree R. Pansare et al, 2011, proposed a system to recognize 26 static hand gestures for ASL

Asian Journal of Convergence in Technology Volume III, Issue II
ISSN No.:2350-1146, I.F-5.11

alphabets from a complex background using the Euclidean distance measure. Experimental setup of the system uses fixed position low-cost web camera with 10 mega pixel resolution mounted on the top of monitor of computer which captures snapshot using Red Green Blue [RGB] color space from fixed distance. This work is divided into four stages such as image preprocessing, region extraction, feature extraction, feature matching. Suryapriya A.K., et al, 2009, presented the development of a frame based approach for speech to sign language machine translation system in the domain of railway and banking. Their work aimed to utilize the capability of Artificial intelligence for the improvement of communication means for physically challenge including deaf-mute people. Input to the system is speech in the domain specified above and the output was a 3D virtual human character playing the sign for the uttered phrases. The system builds up 3D animations from pre-recorded motion capture data. Their work restricted to the spoken Malayalam [15].

III. PROBLEMS IN IMPLEMENTATION HAND GESTURE RECOGNITION

Many basic problems can occur during recognition of static sign language; few of them are listed below:

1. From the literature survey it is observed that there is an inadequate communication and language skill in the majority of deaf children, impacting on poor literature skills in the deaf community.
2. Reality is that deaf schools mainly do not use ISL and nearly 5% of deaf people attend deaf school. The use of ISL is restricted only to vocational programs and short term course.
3. Literature survey shows that ISL was partly influenced by British Sign Language and American Sign Language in the finger spelling system and some other signs, but most are unrelated to European sign system.
4. ISL gesture used in each paper was not common and no specific database available for ISL.
5. As there's hardly any database is available for Devnagri Sign Language (DSL), therefore relative comparison with existing database isn't potential.
6. Many Authors makes the employment of hand gesture gloves that couldn't offers the most effective result for real time gestures.
7. An image has to preprocess and same background is needed before applying to recognition system.

DATA BASE COLLECTION

As there is hardly any database available for Devnagri (Marathi) sign language, this forces us to get new database for DSL Devnagri (Marathi) signing. Thus, so as to get a new database for Devnagri hand gesture pattern tend to use 8 Megapixel webcam. Image capturing can be done by different color space methods such as RGB, Gray and HSV; our system uses the RGB color space model to capture the image. All the static hand gesture images were captured in real time with the resolution of 320 x 240 pixels using USB 2.0 webcam and the data set was prepared. The shot of gesture corresponds to 20 alphabets in Marathi were taken from 25 completely different persons of various group. Total 500 pictures were taken of around picture element size of 320 x 240. Generally, the interval is incredibly high if image size is massive and therefore we tend to reduce the scale of pictures to a good extent making certain no loss of data to be used for signing recognition. So, we have reborn these pictures into a customary size of 284 x 215 pixels. The image capturing process is presented in Figure 1 and the database for class 1 is presented in Figure 2.

Fig.1 Image Capturing

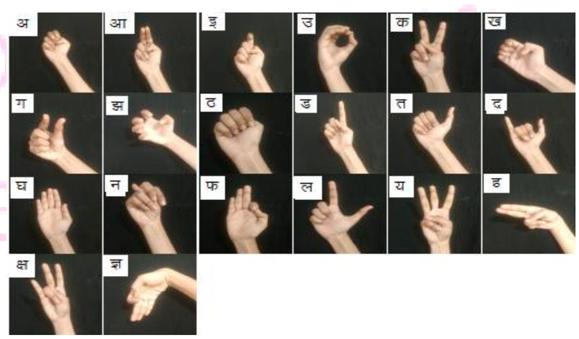


Fig.2 Database for Devnagri gesture for class 1

IV. PROPOSED METHODOLOGY

The scale invariant feature transform (SIFT) algorithm was introduced by Lowe [1] and has been used for feature extraction. This algorithm is one of the most widely used because of the stability over image translation, rotation and scaling and to some extent invariant to change in the illumination and camera viewpoint. They are well localized in both the spatial and frequency domains, reducing the probability of disruption by occlusion, clutter, or noise. Large numbers of features can be extracted from typical images with efficient algorithms. In addition, the features are highly distinctive, which allow a single feature to be correctly matched with high probability against a large database of features, providing a basis for object and scene recognition. The various phases of SIFT algorithm is presented in Figure 3 and input Devnagri gesture to the system is presented in Figure 4.

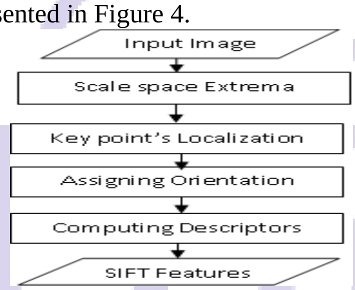


Fig. 3 various phases of SIFT algorithm

Fig. 4 Input images

Following are the major stages of computation used to generate the set of image features:

Scale space extrema detection: The first phase of extreme detection examines the image under various scales to isolate point of picture which is different. These points are called extrema which are potential candidate for image feature.

Key point localization: The next phase, key point location starts from the extrema and select some of these points as a key point. This refinement rejects extrema which is caused by edges of the picture and by low contrast point. Key points are selected on the basis of measures of their stability.

Orientation assignment: The third stage, orientation assignment converts each key point and its neighborhood into a set of vector by computing a magnitude and direction for them. It also identifies other key points that may have been missed in the first two phases. This is done on the basis of key points having a significant magnitude without being an extremism. The algorithm now has identified the final set of key points.

Key point descriptor: The last phase, key point descriptor generation takes a collection of vector in the neighborhood of each key point and consolidates this information into a set of eight vectors called the descriptor. Each descriptor is converted into feature by computing a normalized sum of these vectors.

By applying the above steps, the SIFT algorithm is able to detect 613 key points from the input image as presented in Figure 5 [2],[3],[4],[5].

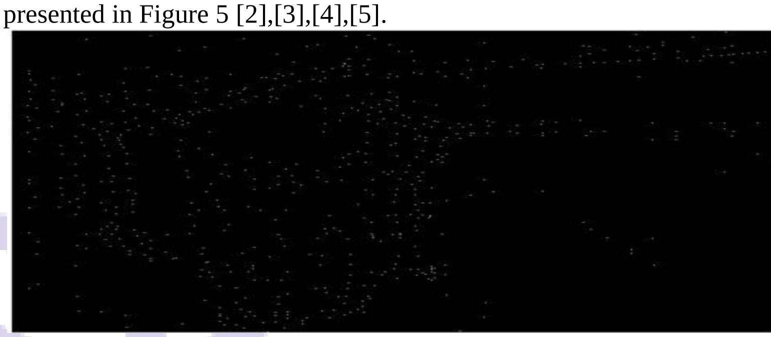


Fig. 5 Total number of Key points (613) for input image

Once the total number of key points extracted from input image, the algorithm is modified to verify the feature matching for intensity variation. Initially the intensity of original image is decreased by 2 and compute the time for various phases of sift algorithm, finding the total number of key points for decreased intensity image and number of key points matched with original image. The same process is applied by decreasing the intensity of reference image viz. 3, 4 and 5 and compare the number of features extracted for original and intensity varying image. The modified algorithm is presented in Figure 6.

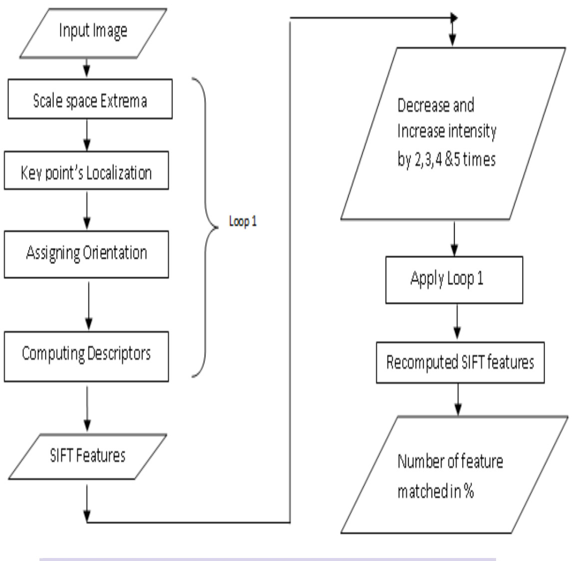


Fig. 6. Various phases of SIFT algorithm with intensity variation

V. RESULT

It can be seen from Figure 6. that the intensity of the reference image is varied by a factor of 2, 3, 4 and 5. Initially, the intensity of the reference image is reduced by various factors and is presented in Figure 7. The total 630 key points are extracted for the image whose intensity is decreased by a factor 5, as presented in Figure 8.

Reference image Intensity decreased by 2 Intensity decreased by 3 Intensity decreased by 4 Intensity decreased by 5

Fig.7 Intensity of reference image is decreased

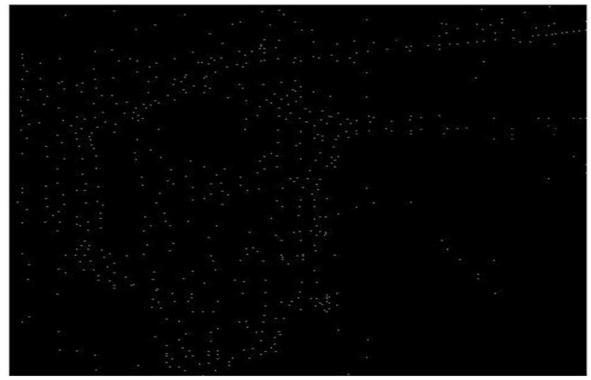
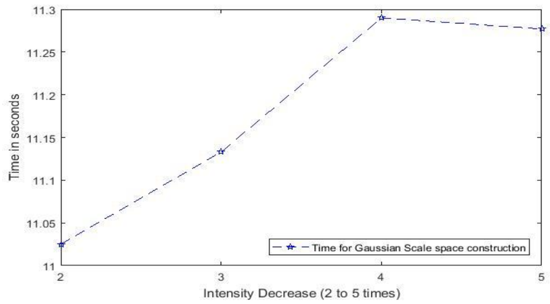


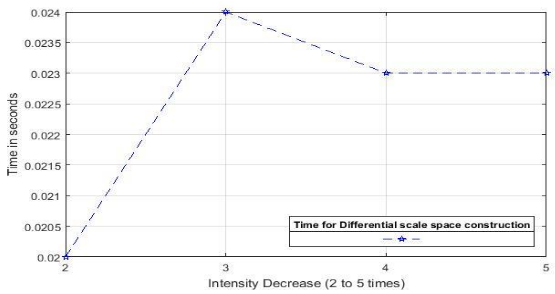
Fig. 8 Key point extraction for Decreased intensity by 5 image (630 key points)

The SIFT algorithm then computes the time for various phases, finding the total number of key points extracted and total key points matched with reference image. The comparative result of SIFT algorithm for reduced intensity is presented in Table 1 and the graphical representation of various phases, total number of key points extracted, total number of key points matched and percentage matching with respect to reference image is presented in Figure 9.

Table 1. Computation time of SIFT algorithm for various intensity values.

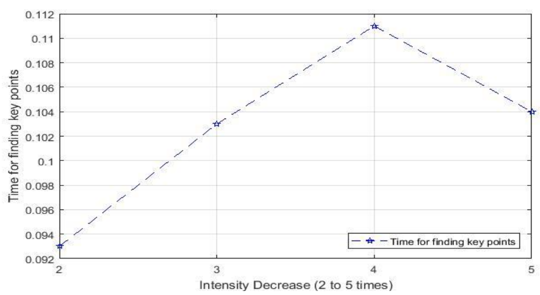
	Computational time for various phases of SIFT algorithm					
Intensity Decreased by	Gaussian Scale space construction	Differential scale space construction	Time for finding key points	Total number of key points extracted	Total number of key points matched	% matching with respect to Reference image
2	11.025	0.020	0.093	613	613	100
3	11.133	0.024	0.103	621	610	99.5106
4	11.290	0.023	0.111	613	613	100
5	11.277	0.023	0.104	630	608	99.1843

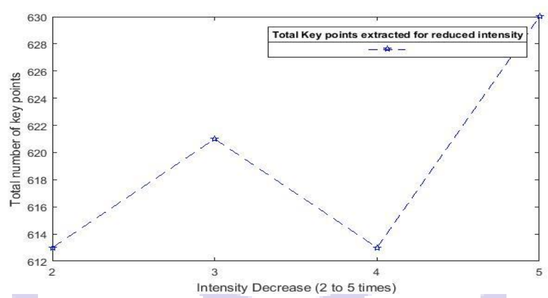




(a)

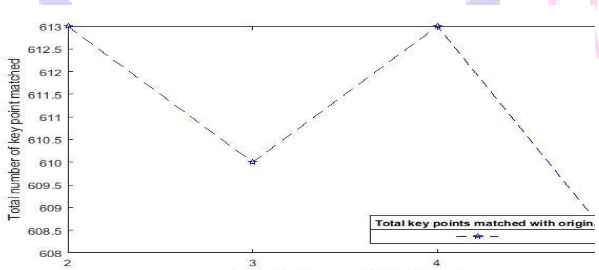
(b)





(c)

(d)



(e)

Fig. 9 Graphical representation of computation time for intensity reduced image (a)Time for Gaussian scale space construction(b)Time for differential scale space construction(c)Time for finding key points(d)Total number of key points extracted(e)Total key points matched.

To determine the matching accuracy, let the count is the number of key points common in two images and kp is the maximum number of key points located in an original image. The percentage matching mp can be calculated by:

$$\% \text{ Matching} = \frac{\text{Count}}{\text{Length (Kp)}} \text{ X } 100$$

(6)

Similarly, the intensity of reference image is increased by a factor 2, 3, 4 and 5, as presented in Figure 10. and total number of key points extracted (630) for increased intensity is presented in Figure 11.

Reference image Intensity increased by 2 Intensity increased by 3 Intensity increased by 4 Intensity increased by 5

Fig. 10 Intensity of original image is increased

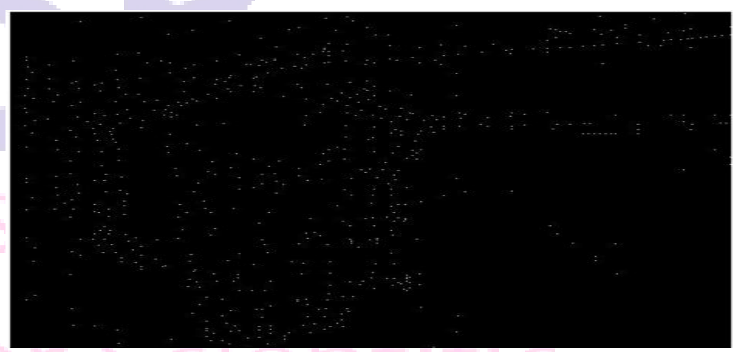


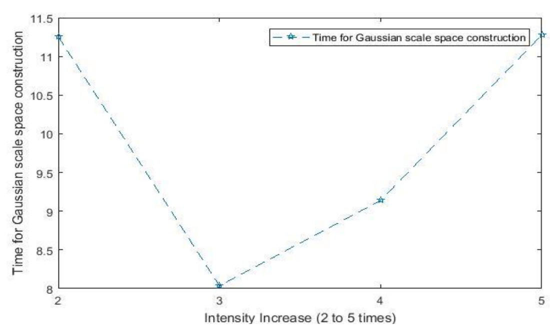
Fig. 11 Key point extraction for increased intensity by 5 image (630 key points)

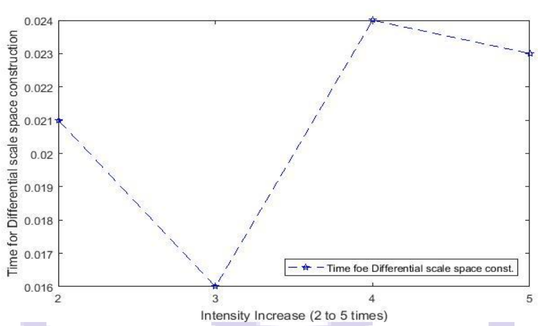
The SIFT algorithm then computes the time for various phases, finding the total number of key points extracted and total key points matched with reference image. The comparative result of SIFT algorithm for increased intensity is presented in Table 2 and the graphical representation of various phases, total number of key points extracted, total number of key points matched and percentage matching with respect to reference image is presented in Figure 12.

Table 2. Computation time of SIFT algorithm for various intensity values.

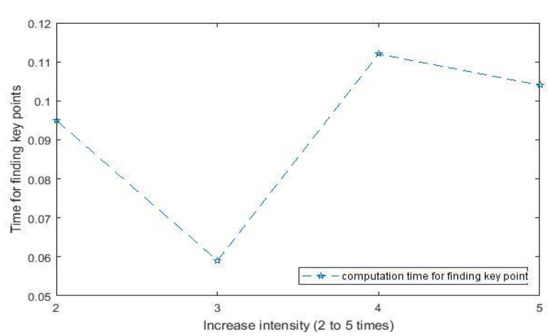
	Computational time for various phases of SIFT algorithm		

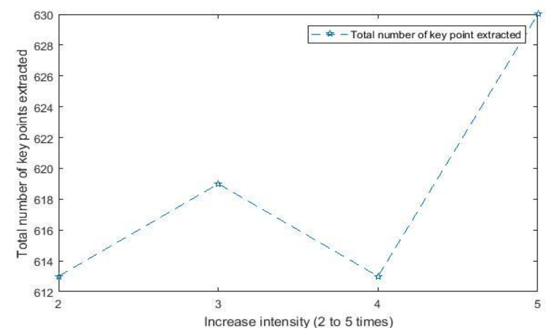
Intensity Increased by	Gaussian Scale space construction	Differential scale space construction	Time for finding key points	Total number of key points extracted	Total number of key points matched	% matching with respect to Reference image
2	11.255	0.021	0.095	613	613	100
3	8.036	0.016	0.059	619	609	99.3475
4	9.139	0.024	0.112	613	613	100
5	11.277	0.023	0.104	630	608	99.1843



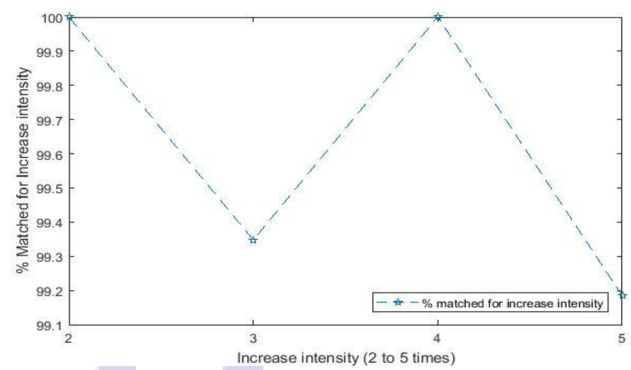


(a)
(b)





(c)
(d)



(e)

Fig. 12 Graphical representation of computation time for intensity increased image (a)Time for Gaussian scale space construction(b)Time for differential scale space construction(c)Time for finding key points(d)Total number of key points extracted(e)Total key points matched.

CONCLUSION

This paper implements the feature extraction and matching process for different intensity values. For the reference image, the time parameter for each phase is computed by SIFT algorithm. The SIFT algorithm is then modified for intensity variation, for increased intensity by a factor of 2, 3, 4, and 5, the algorithm computes the total number of features extracted and matched. The system achieved more than 99% of features matching for increased intensity. Further, the intensity is decreased by the factor of 2, 3, 4, and 5, the algorithm recomputes the time required for various phases and percentage of features matching with respect to the reference image.This implementation shows that the SIFT algorithm provides the reliable matching for different intensity values and hence can be used in the applications where the image is degraded. This paper proposed the system which applicable for recognizing the character of Devnagri (Marathi) sign language. The system can be further modified for the recognition of words and sentence.

REFERENCSE

[1] Lowe. D. G. 2004. *Distinctive image features from scale-invariant keypoints*. International Journal of Computer Vision. 60(2):91-110.

[2] Sandeep Baburao Patil and Dr. Rajesh H. Talwekar, 2017, “ISL Gesture Matching Using Distance Ratio and Rotation Angle” in International Journal of Electronics, Electrical and Computational System IJEECS, ISSN 2348-117X, Volume 6, Issue 7, July 2017, pp-129-136.

[3] Sandeep B. Patil and G.R. Sinha, 2017, “Distinctive feature extraction for Indian Sign Language (ISL) gesture using scale invariant feature Transform (SIFT)” in Journal of the Institution of Engineers (India): series B Springer, Print ISSN-2250-2106 Online ISSN-2250-2114, Vol 98, Issue 1, pp. 19-26.

[4] Sandeep B. Patil and G.R. Sinha, 2015, “Intensity Based distinctive feature extraction and matching using Scale invariant feature transform for Indian Sign Language” in Mathematical Methods and Systems in Science and Engineering, ISBN: 978-1-61804-281-1, Pp-245-255.

[5] Sandeep Baburao Patil and Dr. Rajesh H. Talwekar, 2017, “Implementation of Indian Sign Language Recognition System Using Scale Invariant Feature Transform (SIFT)” in International Journal of Computer Science and Information Security (IJCSIS), ISSN 1947-5500, Vol. 15, No. 2, pp-493-507.

[6] Heera. S.Y., Murthy M. K., and Sravanti V S, 2017, *Talking Hands-An Indian Sign Language to Speech Translating gloves*. International Conference on Innovative Mechanisms for Industry Applications (ICIMIA 2017). Bengaluru. India. IEEE: 746-751.

[7] Singh. A., Sachdeva. R., and Goyal. K. 2017. *Punjabi Sign Recognition and Education system for Differently-able Persons*. International Journal of Advanced Research in Computer Science, University Institute of computer. SBBS University. Punjab. 8(4): 404-406.

[8] Kumar. A., Thankachan. K. and Dominic. M. 2016, *Sign Language Recognition* 3rd InCI Conf. on Recent Advances in Information Technology. Dhanbad. India. IEEE: 422-428

[9] Akoum. A. and Mawla. N. 2015. *Hand Gesture recognition approach for ASL language using Hand extraction algorithm*. Journal of Software Engineering and Applications. 8:419-430

[10] Nanivadekar. P. A. and Kulkarni. V. 2014. *Indian Sign Language Recognition: Database Creation, Hand Tracking and Segmentation*. International Conference on Circuits, Systems, Communication and Information Technology Applications (CSCITA). Broivali West Mumbai India. IEEE: 358-363.

[11] Pol. A. S., Nalbalwar. S. L. and Jadhav. N.S. 2013. *Sign Language Recognition Using Scale Invariant Feature Transform and SVM*. International Journal of Scientific & Engineering Research. 4(6): 1683-1686.

[12] Goyal. S., Sharma. I. and Sharma. S. 2013. *Sign Language Recognition system for deaf and dumb peoples*. International Journal of Engineering Research & Technology. 2(4): 382-387.

[13] Geetha. M. and Manjusha. U. C. 2012, *A Vision Based Recognition of Indian Sign Language Alphabets and Numerals Using B-Spline Approximation*. International Journal on Computer Science and Engineering (IJCSE):406-415.

[14] Kishore. P.V.V. and Kumar. R. P. 2012. *A Video Based Indian Sign Language Recognition System (INSLR) Using Wavelet Transform and Fuzzy Logic*. International Journal of Engineering and Technology. 4(5): 537-542.

[15] Suryapriya. A. K., Suman. S. and Idicula. M. 2009. *Design and development of a frame based MT system for English-to-ISL*. Word congress on nature & biologically inspired computing :1382-1387.