



A SURVEY ON SILENT SOUND TECHNOLOGY USING ELECTROMYOGRAPHY AND IMAGE PROCESSING

MS. Divya P, Shainimol V S

Assistant Professor, MCA Scholar Department of MCA
Nehru College of Engineering and Research centre
divyapakkattil@gmail.com , shainimolvs@gmail.com

Abstract: Everyone has the experience of speaking loudly on the phone during a violent incident while traveling by train or bus. I no longer need to shout for this purpose. It would certainly be a good solution for those who feel frustrated when others speak loudly over the phone. "Silent Sound Technology" aims to recognize all the movements of the lips and convert them into sounds, which can help speechless people to speak, and allow people to make silent calls without disturbing others. Instead of making any noise, your handset will determine the movement of your mouth by measuring muscle activity, and this translates into speech that no one can hear. So, basically, it reads your lips. This new technology will be very useful whenever someone loses voice when speaking or allows people to make silent calls without interrupting others, even we can tell our PIN number to a trusted friend or relative without listening. On the other hand, the listener may hear a clear voice. An amazing feature added to this technology is the "fast polyglot" which can be quickly translated into the language of the user's choice. This translation works for languages such as English, French, and German ". Silent sound technology is an excellent solution for those who have lost their voice but wish to communicate by telephone. This technology allows people to make calls without producing sounds. Developed at Karlsruhe Institute of Technology. This technology basically detects every movement of the lips and internally converts electrical pulses into sound signals and transmits separation to all other surrounding sounds. techniques used to detect trivial sounds, namely electromyography and image processing. This study reviews the future of technological advances that are rapidly changing in the preferred language but, in languages such as different Chinese tones can capture many different meanings. This technology is 99% effective, can be seen on the market for another 5-10 years, and will be used in everyday technology

IndexTerms - Electromyography, muscle movement, electrical pulses, image processing.

I. INTRODUCTION

Silence is the best solution for all situations even with your smartphone. The term mobile phone has become a very popular word in the telecommunications industry. There are many technologies that try to reduce noise pollution and make the environment a better place to live. The new technology is called Silent Sound Technology which will eliminate noise pollution. You are in the theater or in a noisy restaurant or on the bus etc. when there is the noise it becomes a big problem when talking on cell phones. But in the future, this problem will be solved with quiet sounds, a new technology that transforms the movement of the lips into a computer-generated voice on the other side of the phone. This technology aims to monitor the movement of the lips and convert them into computer sounds that can be transmitted over the telephone. So the person on the other side of the phone gets voice information. The concept of "Silent Sound" Technology aims to identify moving lips and transform them into sounds, which can help non-speaking people to speak, and allow people to make silent calls outside without harassing others. In everyone's life, there will be a situation where we just shout and pass a message on the phone because of the loud noise in the background interrupting the people around us and the receiver on the phone and ultimately we all fail in everyone's life there. it has been a situation where we shout just to pass a message on the phone because of the loud noise in the background to convey the message we wanted to convey. To overcome this situation "Silent sound technology" is used. Silent audio technology is a technology that helps you transmit information without using your vocal cords. This technology is intended to provide a solution for those who have lost their voice for some reason and the elderly. This technology is useful if someone wants to pass on a confidential message to a trusted person in a critical situation. It can be used by CID officers in their undercover activities. It can also be used in the field of biomedical. By using this technology the listener can hear a clear voice. By using this technology they

can easily communicate with other people. Technology can also transform you into a fast-paced polyglot. Because electronic pulses are widely available, they can be quickly translated into the language of the user's choice. This technology can be applied to Google Assistant, and Alexa to improve its user experience as Google Assistant does not provide accuracy when used in noisy environments as it picks up inappropriate inputs and provides unexpected output. An important feature of this technology is the translation that works according to the user's choice in English, French, and German.

II. LITERATURE SURVEY

Silent Speech Recognition as an Alternative Communication Device for Persons with Laryngectomy

each year lots of people require surgical elimination of their larynx (voice box) because of trauma or sickness, and thereby require an alternative voice supply or assistive tool to verbally talk. Despite the fact that natural voice is lost after laryngectomy, most muscular tissues controlling speech articulation continue to be intact. Surface electromyographic (SEMG) activity of speech musculature can be recorded from the neck and face and used for automated speech recognition to provide speech-to-textual content or synthesized speech as an alternative manner of conversation. This is genuine even if speech is mouthed or spoken in a silent (subvocal) manner, making it the ideal communication platform after laryngectomy. In this study, eight people a minimum 6 months after total laryngectomy had were recorded using eight SEMG sensors on their face (4) and neck (four) even as analyzing terms made out of a 2,500-word vocabulary. A completely unique set of terms have been used for training phoneme-based reputation fashions for each of the 39 commonly used phonemes in English, and the ultimate terms have been used for testing phrase recognition of the fashions based on phoneme identity from jogging speech. Word blunders quotes were on average 10.3% for the entire 8-sensor set (averaging 9.5% for the pinnacle four individuals), and 13.6% whilst decreasing the sensor set to 4 places according to man or woman (n=7). This takes a look at affords a compelling evidence-of-idea for SEMG-based speech recognition, with the sturdy capacity to further improve recognition overall performance

Smart Assistive Device for Speech Impaired using Silent Sound Technology

One of God's beautiful creations is a people who have the good fortune to have senses such as vision and speech. But some people are unlucky to receive this blessing. People who cannot speak are most affected as they are not able to communicate effectively as is the case with ordinary people verbally. Here people cannot speak using sign language to communicate with others. But some who do not have prior knowledge. This machine can help people with surgeries such as laryngectomy, a laryngectomy that is usually performed in the event of a serious accident or throat cancer. It uses a clever assistant that determines the movement of the lips and converts them into text and sound. The listener on the other hand can be a normal person or a deaf person.

Interactive Silent Speech Recovery:

Jose A. Gonzalez-Lopez, Alejandro Gomez-Alanis, Juan M. Martín Donas, José I. Pérez- Córdoba and the angel M. Gomez interviewed them on the concept of Silent Speech Interfaces for Speech Restoration. Summarizes the research state of the silent speech interface (SSI). SSIs rely on non-acoustic biosignals produced by the human body during speech production to enable communication whenever normal communication is not possible or desirable. In this review, we focus on the first scenario and present the latest SSI study aimed at providing new and alternative communication channels for people with severe speech impairment. SSIs can use biosignals to allow silent communication, such as electrophysiological recording of neural activity, electromyographic (EMG) recording of voice movements, or direct tracking of articulatory movements using imaging techniques. Depending on the disorder, some hearing aids may be more effective than others to capture information related to speech. For example, EMG and imaging techniques are well suited to patients with Laryngectomized, whose voice is almost always perfect but who cannot speak after the vocal cords are removed, but fail in people with severe disabilities. From biological signals, SSIs determine the target message, using automatic speech recognition or speech integration algorithms. Despite major developments in recent years, many modern SSIs are only guaranteed in laboratory settings for healthy users. Therefore, as discussed in this paper, many challenges will still be solved in future research before SSIs can be developed into real-world applications. If these issues can be addressed effectively, future SSIs will improve the lives of people with severe speech impairment by restoring their communication skills.

III. METHODOLOGY

Objectives

The main objectives of system analysis are to find the answers to each business process. It is a process of thinking and involves the creative skills of the System Analyst. It seeks to produce a new and efficient system that meets the current needs of the user and has a wide range of future growth within the organizational challenges. The result of this process is a logical system design. System analysis is a continuous process until a popular and acceptable solution emerges

Proposed Methodologies

The main objectives of system analysis are to find the answers to each business process. It is a thought process and involves the creative skills of the System Analyst. It seeks to produce a new and efficient system that meets the current needs of the user and has a wide range of future growth within the organizational challenges. The result of this process is a logical system design. System analysis is a process that continues until a popular and acceptable solution emerges

Suggested Methods

- Indigenous speakers can speak a sentence silently in their own language, and recipients may hear the sentence translated into their own language. It seems that the native speaker produced the language in a foreign language. Translation technology works in languages such as English, French and German, with the exception of Chinese, where different tones can capture many different meanings.
- Allow people to make silent calls without disturbing others.
- Helping people who have lost their voice due to illness or accident.
- Tell a trusted friend your PIN number by phone without listening to anyone - imagine that no one is reading aloud.
- Silent Sound Techniques are used in the military to communicate private / confidential matters to others.
- This technology is useful for people who have no word
- Telling a PIN or credit card number. the phone is now easier as no one is listening

A. Electromyography

Electromyography is a technique used in quiet sound technology that monitors the small muscle movements that occur when we speak and converts it into electrical pulses that can be converted into speech, without speech. Electromyography (EMG) is a device that can monitor and record the electrical activity produced by the skeletal muscle. EMG is performed using a device called an electromyograph, to produce a record called an electromyogram. The electromyograph detects the electrical activity produced by muscle cells when these cells are activated electrically or emotionally. EMG surface contains

1. Pressure sensor
2. Vibrator Sensor
3. Electric sensor
4. Sense of movement

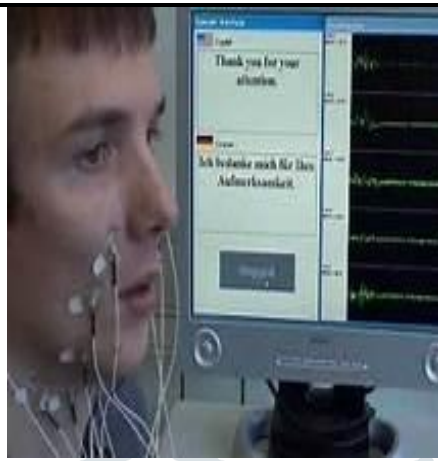


Figure 1: Silent sound translation.

Electromyography consists of a well-drained electrode connected to the face (Fig. 1.1) and an electrical signal generated by the facial muscles is recorded and compared with a recorded signal of the same spoken words. It proved that the electronic signal produces the same of spoken words and the electronic signal can now be transmitted and converted to the same sound signal on the listener's side.

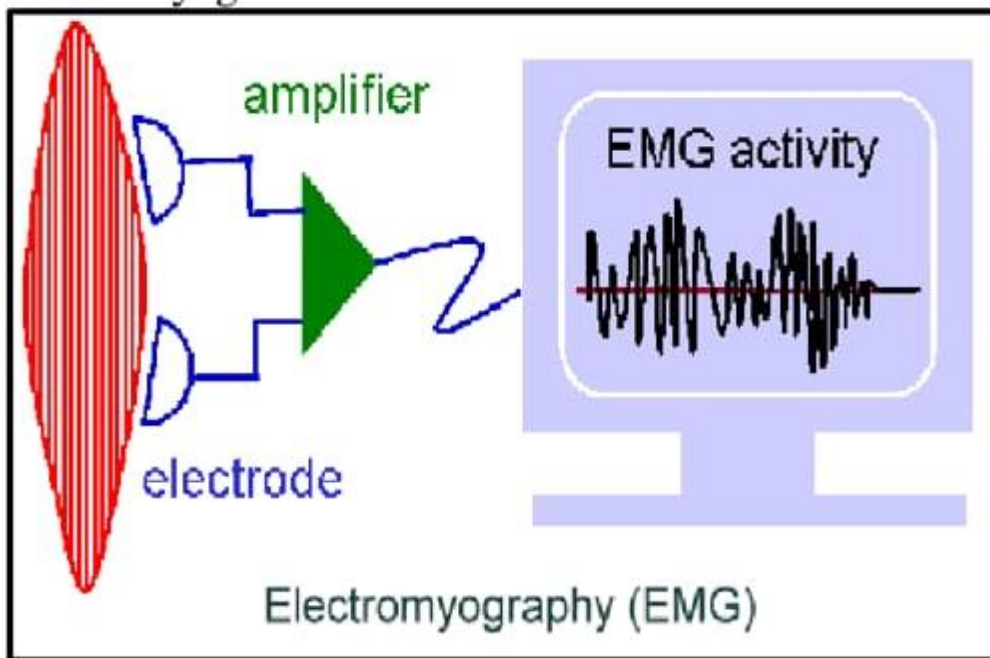


Figure 2: Electromyography

B. Image Processing

An easy-to-use digital image processing method transforms a digital data tape into a film with minimal editing and editing. Then large computer frames are used to process complex data. At this point, the analysts who were to be present were used. In electrical engineering and computer science, image processing is any form of signal processing where input is an image, such as a video frame for a photographer; the image output may be an image or, a set of features or parameters related to the image. Many image processing techniques include treating the image as a two-dimensional signal as well as using standard signal processing techniques. Remote sensor analysis is performed using a variety of image processing techniques and integrated methods

1. Processes analog image

2. Digital image processing

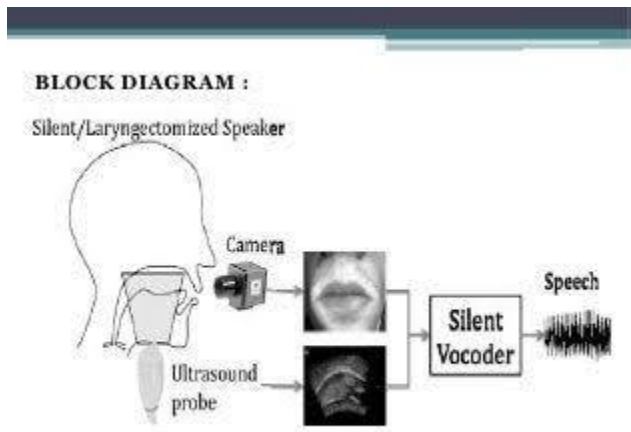


Figure 3: Image processing

1. Analog image processing: -

The analog processing method is used in hard copy data such as photographs or printers. Accepts specific translation features, such as key element, location editing, etc. With so many combinations of remote sensory data it allows us to make a decision not only about what an object is but also its value. Apart from this it incorporates optical photogrammetric techniques that allow accurate measurement of length, width, area, etc. .

2. Digital Image Processing: -

Digital Photo Processing includes a set of techniques for converting digital photos into computers. It contains some errors. In order to overcome errors and omissions to get real data, it needs to take a few steps to process. Digital Image Processing falls into three basic steps:

- 1) Preview Display
- 2) Development
- 3) Release of information.

1) Pre-processing

The initial analysis includes those functions that prepare the data for subsequent analysis that attempts to correct or compensate for system errors. Then the analyst may use the output factor to reduce the data size. Feature extracting is therefore a process of separating the most useful components of data for further research while removing less invasive impacts.

2) Image Enhancement

Improves image interpretation by enhancing brightness between various aspects of Pre-square processing. Development strategies are based on two things

1) Digital data (i.e. with spectral bands and adjustments)

2) Translation Objectives Common enhancements include image reduction, image editing, image enlargement, contrast adjustment, key analysis, text editing and more.

3) Domain Information

In Data Removal Data is subject to quantitative analysis to assign individual pixels to specific classes. Then it is separated. It is necessary to check its accuracy by comparing the categories in the separated images with the known patches on the floor. The end result of this is converted to corresponding signals

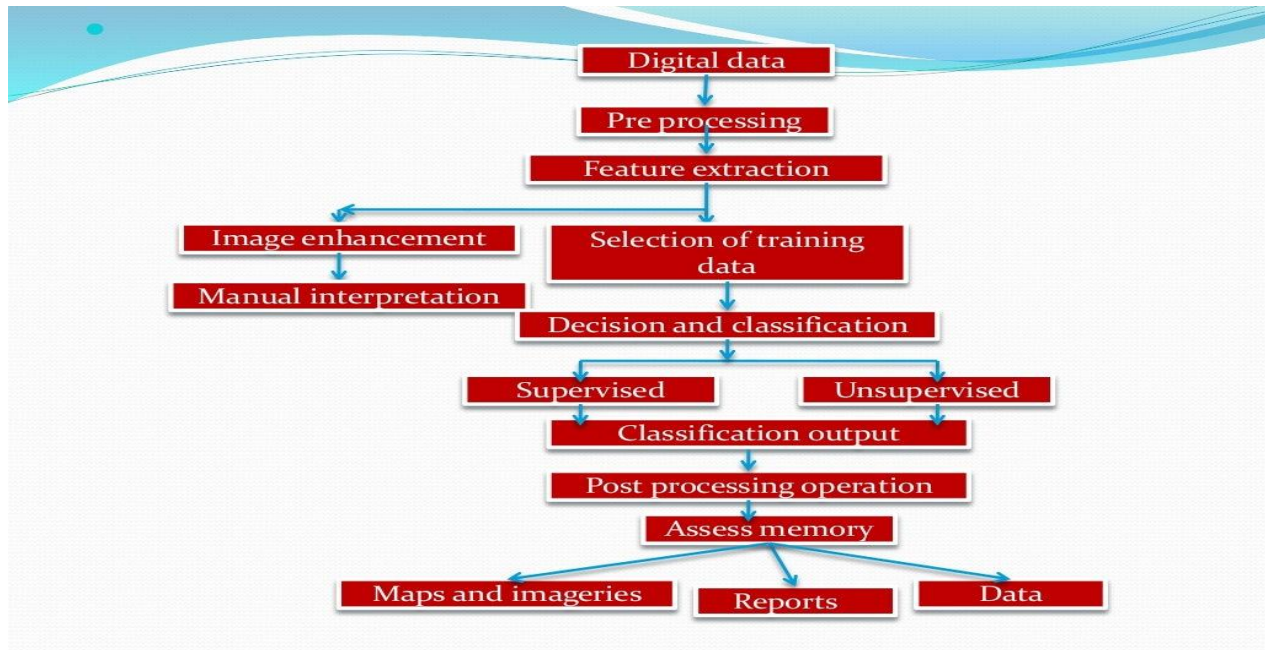


Figure 4: Flow diagram of digital processing

Pre-processing

Initial processing includes those functions that prepare for subsequent analytical data that attempts to correct or compensate for system errors. Then the analyst may use the output factor to reduce the data size. Feature extracting is therefore a process of separating the most useful parts of data for further research while removing less invasive impacts. Reduce the number of variables that should be tested, thus saving time and resources. Initial processing includes those functions that prepare for subsequent analytical data that attempts to correct or compensate for system errors. Digital images are subject to a number of adjustments such as geometry, radiometric, and location,

although all of these methods may not work in all cases. These errors are corrected and can be removed before reaching the user. The investigator must determine which precautionary measures are appropriate on the basis of the type of information that will be extracted from remote data. After the completion of the previous processing, the analyst may use the extraction feature to reduce the data size. Therefore, feature removal is the process of separating the most useful pieces of data to further research while discarding less useful features (errors, noise). Feature removal reduces the number of variables to be tested, thus saving time and resources.

Image Development

Improves image interpretation by enhancing brightness between various aspects of Pre-square processing. Development strategies are based on two things

- 1) Digital data (i.e. with spectral bands and adjustments)
- 2) Translation objectives General enhancements include image reduction, image editing, image enlargement, highlight editing, key component

analysis, text editing and more. Image enhancement activities are performed to improve image interpretation by increasing the contrast between the various elements in the scene. Since the image enhancement method tends to alter real digital data, it is often used for visual (manual) translation and not to analyze additional numbers. Common enhancements include image reduction, image editing, image enlargement, transect output, contrast adjustment, belt measurement, location filtering, Fourier modification, key component analysis, and texture modification.

Extracting Information

In Data Removal Data is subject to quantitative analysis to assign individual pixels to specific classes. Then it is separated. It is necessary to check its accuracy by comparing the categories in the separated images with the known patches on the floor. The end result of this is converted to corresponding signals. Extracting Information is the final step towards the final output of image analysis. After previous processing and image development the remote sensor data is subject to volume analysis to allocate individual pixels to specific classes. . Image classification is based on known and unknown identities to classify the remaining part of an image that includes those unknown pixels. After the classification is complete, it is necessary to check your accuracy by comparing categories in classified images with known patents below. The final result of the analysis contains maps (or images), data and report. These three output components provide the user with complete information about the source data.

Preliminary Processing of Remote Sensitive Images

When remote sensor data is detected by scanning sensors in satellite fields it contains errors and omissions. Preliminary analysis refers to those initial primary analysis activities. Initial processing involves many tasks from the simplest to the most complex and complex. These are classified as

1. Background feature
2. Radiometric correction
3. Geometry Adjustment
4. Location Adjustment

Strategies involved in removing unwanted and disturbing objects such as image / system noise, atmospheric disturbances and sensory movements in image data occur due to dramatic limitations of digital signal encoding, or recording or transmission processes. The removal of these effects from digital data is said to be "restored" to its original or original state, although we do not know what the value is and we must always remember that data acquisition efforts are the only ones that can be introduced. errors. So image restoration involves attempts to correct radiometric and geometric errors.

Enter a domain

The output of the feature does not mean the visual elements that are visible in the image but rather the "statistical" features of the image data such as individual bands or a combination of band values that handle information related to official variations in the scene. Multispectral data is therefore useful in identifying important features of an image. It also reduces the number of spectral bands that need analysis. After the feature release is completed the analyst can work with the required channels or bands, but also each bandwidth with the power of information. Ultimately such processing increases the speed and reduces the cost of analysis.

Image Enhancement Tips

Image enhancement techniques are recommended to make satellite imagery informative and help to achieve the goal of image interpretation. The word enhancement is used to mean to change the appearance of an image in such a way that the information in the image is easily translated into a specific need. The techniques for enhancing the image of the visible red bands and the positive correlation between the visible bands due to the spectral characteristics of the plants are similar as the vegetation increases the appearance of reddish streaks and near-infrared exposure increases. Thus the existence of a correlation between multispectral image bands means that there is duplication of data and that partial analysis is intended to eliminate this termination of Principal Components Analysis (PCA) related to another mathematical method called factor analysis and can be

used for correction. a set of photo bands so that the new bands (called the main parts) are unrelated and arranged according to the variety of images he describes. So the parts are a mathematical summary of the differences that exist in the original band set. To convert real data into primary component axes, conversion coefficients (eigenvalues and eigenvectors) are also used in conjunction with the original pixel values. This line conversion is taken from the covariance matrix of the original data set. These conversion coefficients define the length and direction of the main axis. Such a change is often used as a development function, or prior to data splitting. In the context of PCA, information refers to the variation or distribution of information. Multispectral data is usually less than the number of spectral bands. The purpose of the PCA is to define the magnitude and to adjust the coefficient that determines the set of axis, showing indicators of greater variability. PCA bands are often very interpretive.

Decorrelation Stretch

Key Parts can be expanded and converted to RGB Colors - a process known as stretchable decoration. If the data is converted to the space of the main components and expanded within this space, then the three bands that make up the low-resolution RGB color images will be right angles to one another. In the RGB space, three color segments may be associated, so the expansion effects are not independent of each color. The effect of decorative stretching is usually an improvement in size and filling the space of each color with the remaining unchanged hue. Decorrelation Stretch, as an important component analysis may be based on the covariance matrix or aggregation matrix. The emerging value of the decorative extension is also the natural function of the image used in it.

Canonical Parts

PCA is eligible if a little background information about the incident is available. The analysis of the Canonical component, also called multi-discriminatory analysis, may be useful if information about specific aspects of interest is available. Canonical section axes are available to maximize the variety of user-defined features. Hue, Saturation, and Intensity (HIS) Transform: Hues are produced by mixing red, green, and blue light with red, blue, and colored cubes. The hexane model of hue-saturation-intensity, where the hue is the length of the obtained color control represents an angular position near the top of the hexane, the filling space or purity is given a distance from the center, a straight hexane axis and intensity or number. represented by the maximum hexane width.

Fourier Transform

It works on a single group image. Its purpose is to classify the image by its scale, which is defined as sinusoidal waves with varying amplitudes, waves, and directions. Links to the two-sided space are displayed according to frequency (cycles in each base space). The task of the Fourier Transform is to transform the image of a single band from its local representation to the equal representation of the frequency band and vice versa. The view under the Fourier Transform is that the gray color to form a single band image can be viewed as solid three-dimensional, with lines and columns describing two axes and a grayscale for each pixel giving a third size (z). The Fourier Transform, therefore, provides details of the frequency of each part of the image scale.

IV . RESULT ANALYSIS

Image processing is a physical process used to convert Fourier Transform works on a single group image. Its purpose is to classify the image by its scale, which is defined as sinusoidal waves with varying amplitudes, waves, and directions. Links to the two-sided space are displayed according to frequency (cycles in each base space). The task of the Fourier Transform is to transform the image of a single band from its local representation to the equal representation of the frequency band and vice versa. The view under the Fourier Transform is that the gray color to form a single band image can be viewed as solid three-dimensional, with lines and columns describing two axes and a grayscale for each pixel giving a third size (z). The Fourier Transform, therefore, provides details of the frequency of each part of the image scale. Image processing is a portable process used to convert an image signal into a visual image. The image signal can be digital or analog. The actual result itself can be actual body image or image features. The most common type of image processing is photography. In this process, the image is taken using the camera to create a digital or analog image. In order to produce a visual image, the image is processed using appropriate technology based on the type of input source. Digital

Photo Processing involves a set of techniques for converting digital images to computers. Digital image processing is the use of computer algorithms to perform image processing in digital images. As a sub-category or digital signal processing field, digital image processing has many advantages over analog image processing. It allows a much wider range of algorithms to be used in input data and can avoid problems such as sound formation and signal distortion during processing. As images are defined in two dimensions (perhaps more) digital image processing can be modeled in the form of Multidimensional Systems. Analogue processing techniques used in hard copy data such as images or printers. Image analysis in visual techniques uses certain translation features. Muscle tissues at rest often do not function electronically. After the electrical activity caused by the irritation of the needle is reduced, the electromyograph should not detect spontaneous activity (i.e., the resting muscle should be electrically silent, except in the area of neuromuscular junction, i.e., under normal conditions., It works automatically). When the muscle is voluntarily shortened, the force of action begins to emerge. As muscle contraction capacity increases, more muscle fibers produce energy to do something. Once the muscle is fully absorbed, there should be an ineffective group of action of various levels and amplitudes (complete rental pattern and distraction) Calls' - calls that are not answered or unknown because the user is in a situation where he or she cannot speak - not just in business meetings, but in everyday situations. According to research, these 'lost calls' cost \$ 20 billion a year worldwide. For the mobile operator, these are the potential benefits currently left on the table. If these 'lost calls' are answered and can be held without making noise, there is a high potential for increased profits.

IV. CONCLUSION

Silent Sound technology aims to detect all moving lips and turn them into sounds, which can help speechless people to speak, and allow people to make silent calls without disturbing others. Instead of making any noise, the cell phone can detect the movement of the mouth by measuring muscle activity and converting this into speech that the other person can hear. So it basically reads lips. It will be one of the most innovative and useful technologies and in the near future these technologies will be used in everyday life. the analysis is ongoing, the technology to be used in the Associate in Nursing workplace. So Silent Sound Technology, one of the latest trends in the field of information technology uses —Speaking without speaking!. It will be one of the new and useful technologies and in the future these technologies will be used in our daily lives. Silent Sound technology aims to detect all the movement of the lips and convert them into sounds, which can help speechless people to speak, and allow people to make silent calls without disturbing others. Instead of making any noise, your handset will determine the movement of your mouth by measuring muscle activity, and this translates into speech that no one can hear. So, in fact, it reads your lips With a quiet voice technology will be installed on cell phones or headphones where the earphones will detect the movement of the lips and jaws and gain electrical energy that will be converted into a sound signal before being transmitted by developers. they say the device is 99 percent effective. Silent Sound Technology. these technologies will be used in our daily lives. In the future, this technology will be used in everyday life. So Silent Sound Technology is one of the latest trends in the field of data technology.

VI . REFERENCES

- [1] Denby, B., Schultz, T., Honda, K., Hueber, T., Gilbert, J.M., Brumberg, J.S., Silent Speech Interfaces, Speech Communication Volume 52, Issue 4, 270-287, April 2010.
- [2] Hueber T, Benaroya E-L, Chollet G, Denby B, Dreyfus G, Stone M. (2010). Development of a silent speech interface driven by ultrasound and optical images of the tongue and lips. Speech Communication, Volume 52, Issue 4, 288–300, April 2010.
- [3] <http://www.dellchallenge.org/projects/silent-sound-technology>
- [4]<http://www.telecomspace.com/content/cebit-2010-silent-sound-technology-endless-possibilities>
- [5] <http://en.wikipedia.org/wiki/Electromyography>
- [6] Nigg B.M., & Herzog W., 1999. Biomechanics of the Musculo-Skeletal system, Wiley, Page 349.

[7]<http://www.techpark.net/2010/03/04/silent-sound-technology-an-end-to-noisy-communications/>

[8]<http://www.telecomspace.com/content/cebit-2010-silent-sound-technology-endless-possibilities>

[10]<http://www.techpark.net/2010/03/04/silent-soundtechnology-an-end-to-noisy-communications/>

[11] Shehjar Safaya , Kameshwar Sharma , Silent Sound Technology- An End to Noisy Communication, Speech Communication, Vol. 1, Issue 9, November 2013

[12] www.dellchallenge.org.

[13]Anish Kumar , Rakesh Raushan , Saurabh Aditya , Vishal Kumar Jaiswal , Mrs. Divyashree, An Innovative Communication System For Deaf, Dumb and Blind People Volume 5 Issue VI,June 2017 IC Value: 45.98 ISSN: 2321-9653

[14]S D, Lalitha & K K, Thyagarajan. (2016). A Study on Lip Localization Techniques used for Lip reading from a Video. International Journal of Applied Engineering Research. 11. 611- 615.

[15]Hueber, T., Benaroya, E-L., Chollet, G., Denby, B., Dreyfus, G., Stone, M., Development of a Silent Speech Interface Driven by Ultrasound and Optical Images of the Tongue and Lips, Speech Communication (2009)

[16]J. Chung, A. Senior, O. Vinyals and A. Zisserman, "Lip Reading Sentences in the Wild," in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017 pp. 3444-3453

[10]<http://www.techpark.net/2010/03/04/silent-soundtechnology-an-end-to-noisy-communications/>

[11] Shehjar Safaya , Kameshwar Sharma , Silent Sound Technology- An End to Noisy Communication, Speech Communication, Vol. 1, Issue 9, November 2013

[12] www.dellchallenge.org.

[13]Anish Kumar , Rakesh Raushan , Saurabh Aditya , Vishal Kumar Jaiswal , Mrs. Divyashree, An Innovative Communication System For Deaf, Dumb and Blind People Volume 5 Issue VI,June 2017 IC Value: 45.98 ISSN: 2321-9653

[14]S D, Lalitha & K K, Thyagarajan. (2016). A Study on Lip Localization Techniques used for Lip reading from a Video. International Journal of Applied Engineering Research. 11. 611- 615.

[15]Hueber, T., Benaroya, E-L., Chollet, G., Denby, B., Dreyfus, G., Stone, M., Development of a Silent Speech Interface Driven by Ultrasound and Optical Images of the Tongue and Lips, Speech Communication (2009)

[16]J. Chung, A. Senior, O. Vinyals and A. Zisserman, "Lip Reading Sentences in the Wild," in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017 pp. 3444-3453.