LATVIAN JOURNAL OF PHYSICS
AND TECHNICAL SCIENCES
2022, N 2

DOI: 10.2478/lpts-2022-0007

MAXIMUM SHANNON INFORMATION DELIVERED IN A LECTURE

L. Alksne[1], A. Ozols [2]

[1]Liepaja University, 14 Liela Str., Liepaja, LV-3401, LATVIA
[2] Faculty of Electronics and Telecommunications,
Riga Technical University, 12 Azenes Str., Riga, LV-1007, LATVIA
[1] linda.alksne@liepu.lv
[2]andris.ozols@rtu.lv

The aim of our paper is to evaluate the maximum Shannon (syntactic) information carried through a video lecture. To achieve the aim, we have considered a natural lecture delivered by a lecturer as a signal transmitted over the physical communication channel consisting of a sound sub-channel and light sub-channel. Receivers are eyes and ears of listeners whose physical characteristics are taken into account. The physiological, neurological and cognitive aspects of the problem are neglected in calculations. The method has been developed to calculate the absolute maximum values of Shannon information characteristics of a natural lecture basing on the capacity formula of continuous communication channel and physical considerations taken into account for the first time, to our knowledge. Maximum Shannon information characteristics (entropies of sound and light frames, amounts of total acoustical and optical information, capacities of sound and light sub-channels, total amount of information and total capacity) of a natural lecture perceived by the audience have been calculated. These values are the upper bounds of a video lecture. The obtained results are discussed in the paper. After some modification, the proposed method can be practically applied for the optimization of both natural and video lectures because there is some correlation between syntactic and semantic information characteristics.

Keywords: *Channel capacity, communication channel, entropy, lecture, semantic information, syntactic information, video lecture.*

1. INTRODUCTION

Conceptually, information can be thought of as being stored in or transmitted as variables that can take on different values. Informally, we obtain information from a variable by looking at its value, just as we receive information from an email by reading its contents. In the case of the variable, the information is about the process behind the variable [1]. Generally, there are two kinds of information, which are important with respect to a lecture. Shannon information provides the so-called syntactic information, which reflects the amount of statistical correlation between systems. By contrast, semantic information refers to information, which is in some sense meaningful for a system, rather than merely statistical [2]. The aim of the paper is to calculate the maximum syntactic Shannon information delivered in a video lecture. To achieve the aim, the delivery process of a natural lecture is considered from the physical point of view. Further in the paper, we always mean the Shannon information and the quantities related to it, if not mentioned otherwise.

Shannon entropy of the message ensemble X consisting of messages x_i is the average amount of information $I(x_i)$ contained in a message:

$$H(X) = \sum_{i=1}^{n} p(x_i)\, I(x_i) = \sum_{i=1}^{n} p(x_i) \log_b \frac{1}{p(x_i)} = -\sum_{i=1}^{n} p(x_i) \log_b p(x_i). \tag{1}$$

It depends only on the statistic nature of the message source expressed by message probabilities $p(x_i)$. Here b=2 corresponds to the binary logarithm, and n is the number of messages [3].

Different applications of Shannon entropy are used to evaluate the information content of videos. As a rule, this entropy (or quantities related to it, such as mutual information, joint entropy, conditional entropy) is calculated by computer programs (e.g., [4]) within each 2D video frame based on intensity and colour of frame elements. However, entropy there is in relative units and serves to evaluate the changes. Analogously, the speech entropy is calculated and used [5].

Based on the differences between video frames, the entropy is able to serve as a measure of the complexity of changes. Due to content dependency, however, the relative entropy changes in the sequence of video frames, being a better indicator for detection [6].

Experimental video results of Xuguang Zhang [7] have shown that the panic crowd motion state has higher entropy, and the normal crowd state has lower entropy. When a panic behaviour of a crowd occurs, the pedestrians often move hurry-scurry. As the pedestrians are moving, the attributes (such as gender and age) of pedestrians are different. The speeds of the movements of individuals are also different. The motion information of the body parts (arms, torso and legs) of an individual is also different. The motion flow of the crowd video represents a state of disorder. In 2016, Luo also proposed a detection algorithm based on skeleton entropy by using the information from RGBD videos [8]. The entropy is analysed in terms of the angles of the body skeleton to find whether the values of the information entropy are significantly higher in abnormal videos than in normal videos in order to detect the most abnormal behaviours, such as fight, robbery and chaos. The principle of the highest frame entropy changes is used in video advertising [9]. The joint entropy changes and conditional entropy changes of *x* and *y* coordinates are used for the classification of landscapes [10]. Černekova et al. have demonstrated that mutual entropy and joint entropy between the frames can be used to detect cuts and extract key frames [11].

There are also other applications of Shannon entropy, for example, in video compression [12] and security [13]. An important achievement is Memorability-Entropy-based video summarization. Authors predict the memorability score by using the fine-tuned deep network and calculate the entropy value of the images. The frame with the maximum memorability score and entropy value in each shot is selected to constitute the video summary. Memorability is the quality of being worth remembering [14].

However, the authors are mostly interested in entropy and related information characteristics to improve the efficiency of video lectures. Thus, entropy can be calculated for random time slot of a random video lecture captured in classroom and for the same length of video lecture followed by a set of rules on how to produce a good video lecture [15]. These rules have already given some results while transmitting information to people. Prior studies have investigated the effect of instructional video on learning outcomes [16]. Thus, it is empirically shown that there is some correlation between Shannon entropy of videos and the efficiency of video lecture, i.e., the semantic information content.

For further development, it would be important to have not only relative information characteristics of videos, but also absolute ones. As far as we know, there are no papers dealing with the absolute Shannon information carried by lectures or by natural scenes. Therefore, the method of expert evaluations of the quality of lectures is used [17]. Basing on Shannon's communication theory (see Section 2), we propose a method how to calculate Shannon information delivered in a natural lecture using only physical properties of human sensors (eyes and ears) and neglecting any physiological processes in the brain, which are too complicated to be taken into account in this first calculation attempt. We quantitatively evaluate the maximum Shannon (syntactic) information, which can be physically delivered in a lecture of a certain length and perceived by the audience. As mentioned above, syntactic information refers only to the quantity of unexpected data not to their meaning. We suppose that a lecturer speaks and shows slides and demonstrations for a certain period of time. Thus, the audience receives a certain amount of optical and acoustical information by means of eyes and ears. This is the maximum possible information. If the lecture is captured by the video with sound and later reproduced, the information delivered in this video lecture will be reduced because of technical limitations of video recorder (limited optical and acoustical bandwidths, etc.).

In this article, natural lecture delivered by a lecturer is treated as a noiseless communication channel consisting of a sound sub-channel and light sub-channel. Maximum transmitted total amount of information in both sub-channels and in the whole channel is calculated as well as the corresponding channel capacities. Further mathematical and experimental development of the presented approach can be applied for the optimization of both natural and video lectures.

2. METHOD

The idea to calculate the Shannon information of a lecture is based on the capacity formula of continuous communication channel:

$$C_{sample} = \max I(X,Y) \text{ bits/sample,} \quad (2)$$

where $I(X,Y)$ is an averaged mutual information of ensemble of received signals Y with respect to ensemble of transmitted signals X per sample[3, p.586]. If the channel has a fixed bandwidth B, then the maximum signalling rate is Nyquist rate $2B$ according to the sampling theorem, and channel capacity is as follows:

$$C = 2B \max I(X,Y), \text{ bits/s.} \quad (3)$$

Transmission of a lecture to the audience can be regarded as a communication channel consisting of a sound sub-channel and a light sub-channel. They have different known frequency bandwidths determining the maximum signalling rates. The main problem is to determine the mutual information of sound sub-channel and light sub-channel. However, we can avoid this problem neglecting the noise. In this case, mutual information in each sub-channel can be replaced by its Shannon entropy $H(X)$ [3]. However, for continuous signals the probabilities of their values x $p(x)=w(x)dx$ in Eq. (1) become infinitesimally small (here $w(x)$ is the differential probability distribution function). To overcome this problem we assume that we can approximately replace infinitesimally small probabilities $p(x)=w(x)dx$ by small but finite probabilities $p(x)=w(x)\Delta x$, which can be numerated. Shannon entropies can be calculated by appropriate quantization of different parameters x coding the transmitted information (e.g., sound and light intensity or frequency) in intervals Δx. In this article, we are interested in the maximum Shannon information; therefore, quantization intervals are determined based on the human resolving power of coding parameters. Thus, based on this hypothesis, we can calculate not only the maximum Shannon entropies of sound and light channels but also maximum capacities of these channels and transmitted information during the lecture.

Further we shall explain our method in more detail using as an example the woman lecturer who delivers a lecture, the duration of which is ten minutes (t = 10 min).

3. DETAILED EXPLANATION OF THE METHOD RESULTS

At first, we consider sound and light sub-channels separately.

3.1. Entropy of the Sound Channel

We assume that the lecturer is a woman whose voice is soprano. Her speech is characterised by sound intensity, I_s, sound intensity modulation frequency bandwidth, $F_{s\ \mathrm{mod}}$, and by sound frequency, f_s. Fundamental frequencies of soprano belong to the spectral interval f_s = 260–1050 Hz [18]. They bring the main sound energy of the speaker. In the case of soprano singer harmonics up to the fourth can also play a significant role [19]; however, our woman is a lecturer, not a singer. Thus, we assume that her sound frequency bandwidth is F_s = 790 Hz. Humans can perceive sound intensity changes with the modulation bandwidth of about $F_{s\,mod}$ = 500 Hz [20]. This means that we can hear faster than we see because the corresponding light intensity modulation bandwidth for humans is only about 10 Hz (see Section 3.3). It should be stressed that both intensity bandwidths refer to the envelope function of intensity time dependence.

Due to the limited sound intensity modulation bandwidth, the sound signal can be sampled with the sampling time interval [3]:

$$\Delta t_s = \frac{1}{2 \times F_{s\,mod}} = 1 \times 10^{-3} s \quad ,$$

and we can consider only sampled moments instead of the continous time. At each sampled time moment, the listeners perceive the sound frame – the sound spectrum $I_s(f_s)$, which carries the spectral sound information at the given moment (this is necessary because only the envelope of sound intensity time dependence is taken into account for sampling). In principle, there is an infinite number of possible sound frames at the certain sampled moment of time because sound intensity and frequency vary continously. As mentioned above in Section 2, we have to make these variables discrete. On the other hand, this quantization occurs naturally because the human ear has limited intensity and frequency resolution power. Thus, a sound frame is a 2D message from the point of view of communication theory. The lecturer is a source of these messages, which follow at frequency $2F_{s\,mod}$. The average information per such a message by definition equals the Shannon entropy of the ensemble of all such messages according to Eq. (1).

Thus, the calculation of the Shannon entropy is based on the making discrete the sound frame variables, I_s and f_s, taking into account the resolution power of human ear. We divide the sound frequency in intervals Δf_s =1 Hz [the minimum frequency difference which can be resolved by ear [21]] and sound intensity intervals $\Delta I_s = 1.2 \times 10^{-6}$ W/m^2 corresponding to the minimum noticable sound pressure changes of 0.5 dB [22]. Thus, there are

$$\frac{F_s}{\Delta f_s} = \frac{790}{1} = 790$$

discrete sound frequency intervals. If we assume that the highest sound intensity of the speaker corresponds to fortissimo (100 dB or 10^{-2} W/m^2), then (keeping only integers because fractional sound intensity interval cannot be resolved by a human ear) we have

$$\frac{I_{s\max}}{\Delta I_s} = \frac{10^{-2}}{1.2 \times 10^{-6}} \approx 8333$$

sound intensity intervals which can be resolved by the ear. In this way, we can divide the sound frame into $N_s = 790 \times 8333 =$ 6583070 2D ($\Delta I_s \times \Delta f_s$) cells.

Depending on the sound signal, each cell can be either filled or empty. For simplicity we assume that the fulfilment probabilities are equal for all cells. Sound frames differ by the number and distribution of filled and empty cells. As mentioned above, each frame can be regarded as a message. To get the maximum frame entropy the probabilities of these messages should be equal [3]. This can be readily seen from the Shannon entropy expression (1) letting all the probabilities of n messages to be equal:

$$p(x_i) = 1/n\,, \qquad (4)$$

$$H(X) = \log_2 n. \qquad (5)$$

How many messages are there? For the fixed number of filled cells, k, there are $C_{Ns}{}^{k}$ messages, where $C_{Ns}{}^{k}$ is the number of combinations for fixed k of a set of N_s elements. Number k can vary from 0 to N_s. Therefore, we have to sum all $C_{Ns}{}^{k}$ for $k = 0, 1,2,..... N_s$ and to obtain the number of combinations for all k. The result is [23]:

$$\sum_{k=0}^{N_s} C_{N_s}^{k} = 2^{N_s}. \tag{6}$$

Thus, the number of messages is $n = 2^{Ns}$ and the maximum entropy of the sound frame is [3]:

$$H_{s\ max} = \log_2 2^{Ns} = N_s. \tag{7}$$

Previously, we have found that N_s = 6583070; therefore, $H_{s\ max}$ = 6583070 bits.

3.2. Total Maximum Information in the Sound Sub-channel and its Capacity

Let us find the total maximum information transmitted over the sound sub-channel. During the lecture of duration t = 600s, the number of transmitted sound frames is:

$$\frac{t}{\Delta t_s} = \frac{600}{1\times10^{-3}} = 600000.$$

Each sound frame carries the maximum average information $H_{s\ max}$ = 6583070 bits as shown in the previous section. Assuming that all messages (sound frames) are statistically independent, we find that the total maximum transmitted information over the sound sub-channel is [3]:

$$Info_{s\ max} = \left[\frac{t}{\Delta t_s}\right] \times H_{s\ max}. \tag{8}$$

Putting all the known quantities t=600s, $\Delta t_s = 1.0\times10^{-3}$ s and $H_{s\ max}$ = 6583070 bits in Eq.(8) yields $Info_{s\ max} = 3.95\times10^{12}$ bits. Here and further we keep the number of digits to have accuracy not worse than 0.3 %.

The communication channel capacity is defined as the maximum amount of information transferred per second [3]. Thus, the maximum capacity of our noiseless sound sub-channel is:

$$C_{s\ max} = \frac{Info_{s\ max}}{t} = \frac{H_{s\ max}}{\Delta t_s}. \tag{9}$$

Putting the above-mentioned values of $Info_{s\ max}$ and t, or of $H_{s\ max}$ and It_s in Eq. (9), we get $C_{s\ max} = 6.58\times10^{9}$ bits/s.

3.3. Entropy of the Light Frame

We consider the optical information transmission in light sub-channel analogously to that of sound sub-channel. Light sub-channel is characterized by the light intensity $I_{l,}$, light intensity modulation frequency bandwidth $F_{l\,mod}$ describing the speed of the light intensity temporal changes, light frequency f_l determining the image colours (we neglect the eye sensitivity spectral dependence), field of view angles along the transversal coordinates of the scene x and y, θ_x and θ_y, and the transmission time t, which is the same as for the sound sub-channel.

The maximum light frame (image) rate is determined by the light intensity modulation bandwidth of eye, $F_{l\,mod,}$ which is equal to 10 Hz at the eye contrast sensitivity function level of 200. This level corresponds to the light intensity resolution of 0.5 % by the human eye. Smaller intensity differences cannot be resolved [24]. This means that there are

$$\frac{I_{l\max}}{\Delta I_l} = \frac{1}{0.005} = 200$$

discrete detectable light intensity intervals. On the other hand, according to the sampling theorem, light intensity envelope time dependence can be sampled with the sampling time interval

$$\Delta t_i = \frac{1}{2\times F_{l\max}} = \frac{1}{2\times 10Hz} = 5\times 10^{-2} s\,.$$

For fixed sampling time, the audience perceives optical image, the light frame, determined by the light intensity I_l dependence on f_l, θ_x and θ_y. This image is a 4D message from the point of view of Shannon's information theory. Again, the average information per such a message by definition equals to the Shannon entropy of the ensemble of all such messages, as shown in Eq. (1).

Further, light frame entropy calculation is analogous to the sound frame entropy calculation described in Section 3.1. We divide the light frequency into intervals Δf_l= 3.0×10^{12} Hz. This is the minimum frequency difference which can be resolved by eye [25]. The visible spectral range in video lecture we assume to be from 400 to 750 nm [26] which corresponds to the light frequency range F_l = 3.5×10^{14} Hz. Thus, there are (keeping again only integers)

$$\frac{F_l}{\Delta f_l} = \frac{3.5\times 10^{14}}{3.0\times 10^{12}} = 116$$

resolvable light frequency intervals. [Here we would like to stress that the number of colours perceived by a human is much larger. Human brain constructs colours from the perceived frequencies of different intensities within these intervals. We are restricting ourselves only to physical rather than physiological processes involved in the delivery of lecture.]

The minimum intervals of viewing angles, θ_x and θ_y, are determined by the accepted minimum detectable light intensity changes of 0.5% and are equal to $\Delta\theta$=15 arc minutes=4.36×10^{-3} rad [24]. The maximum horizontal eye field of view along the horizontal x-axes is $\theta_{x\ max}$ = 160^0 = 2.79 rad and along the vertical one θ_{ymax} = 130^0=2.27 rad [18], [27]. Therefore, there are

$$\frac{\theta_{x\max}}{\Delta\theta} = \frac{2.79}{4.36\times 10^{-3}} = 640$$

and

$$\frac{\theta_{y\max}}{\Delta\theta} = \frac{2.27}{4.36\times 10^{-3}} = 520$$

intervals along the θ_x and θ_y axes, respectively (analogously as in the sound channel, fractional angular intervals cannot be resolved by a human eye).

There are N_l 4D cells in one light frame, where

$$N_i = \frac{I_{l\max}}{\Delta I_l}\times\frac{F_l}{\Delta f_i}\times\frac{\theta_{x\max}}{\Delta\theta}\times\frac{\theta_{y\max}}{\Delta\theta} = 7720960000 \approx 7.72\times 10^9\,.$$

We can directly apply Eqs. (6) and (7) to the light sub-channel replacing the number of 2D cells N_s by the number of 4D cells N_l, because the result depends only on the number of cells but not on their dimensionality. Thus,

$$H_{l\ max} = \log_2 2^{N_l} = N_l \tag{10}$$

and $H_{l\ max}$ = 7.72×10^9 bits.

3.4. Total Maximum Information in the Light Sub-channel and its Capacity

Total maximum information in the light sub-channel and its corresponding capacity are calculated in the same way as in the case of sound sub-channel (Section 3.2). Therefore,

$$Info_{l\ max} = \left[\frac{t}{\Delta t}\right] \times H_{l\ max} \quad (11)$$

and

$$Info_{l\max} = \frac{600}{5\times10^{-2}} \times 7.72\times10^{9} \approx 9.26\times10^{13}\,bits\ .$$

Instead of (8) in the light sub-channel we have

$$C_{l\ max} = \frac{Info_{l\ max}}{t} = \frac{H_{l\ max}}{\Delta t_l} \quad (12)$$

and, consequently, putting the values $Info_{l\ max}$ and t, or $H_{l\ max}$ and Δt_l in Eq.(10) we get $C_{l\ max} \approx 1.54\times10^{11}$ bit/s.

3.5. Total Maximum Information of the Lecture and the Capacity of the lLcture as a Communication Channel

The obtained information characteristics of sound and light sub-channels enable one to find the maximum information and maximum capacity of the whole lecture by simply summing the corresponding quantities because we can assume them to be independent. In this way, the maximum information delivered by the considered lecture is as follows:

$$Info_{max} = Info_{s\ max} + Info_{l\ max}, \quad (13)$$

and the maximum capacity of the whole lecture channel is as follows:

$$C_{max} = C_{s\ max} + C_{l\ max}\ , \quad (14)$$

because amounts of information can be always summed, and the information transmittance time is the same for both sub-channels. Putting the corresponding quantities in Eqs.(13) and (14) we get $Info_{max} = (3.95\times10^{12} + 9.26\times10^{13})$ bits $\approx 9.65\times10^{13}$ bits, and $C_{max} = (6.58\times10^{9} + 1.54\times10^{11})$ bits/s $\approx 1.61\times10^{11}$ bits/s.

4. DISCUSSION

It is clearly seen from results of the previous section that Shannon (syntactic) information characteristics of the whole lecture channel are almost completely determined by the light sub-channel because information carried by the sound sub-channel is lower by more than one order of magnitude. The ratio of entropies of light and sound frames is even higher:

$$\frac{H_{l\max}}{H_{s\max}} = \frac{7.72\times10^{9}}{6583070} \approx 1173\ .$$

Thus, the contribution of the lecturer's voice to the syntactic information is almost negligible. At the first glance, this result seems to be expected because vision ranks highest in the hierarchy of human senses. In the lecture, it seems that only the demonstration of slides is necessary. On the other hand, this is a paradoxical result because practically we know that the role of the lecturer is of primary importance. This paradoxical result is the consequence of neglecting the meaning of the lecture when calculating the syntactic Shannon information. Not only the voice but also the intonation and gestures of the lecturer play an important role expressing the attitude of the lecturer to the content. It should be also noted that if the lecturer used additional sound accompaniment, e.g., music, with larger sound frequency bandwidth up to 20 kHz (the maximum bandwidth of a human ear [18]), then $H_{s\ max}$, $Info_{s\ max}$ and $C_{s\ max}$ would increase by more than one order of magnitude achieving their maximum possible values comparable to the corresponding light sub-channel parameters. This situation takes place in concerts.

We can compare the above calculated capacities of sound and light sub-channels with the known information capacities of the human hearing and human sight. We have found that $C_{s\ max} = 6.58\text{x}10^9$ bits/s and $C_{l\ max} = 1.54\text{x}10^{11}$ bits/s whereas the capacity of human hearing channel is about 10^4 bits/s and human sight channel is about 10^7 bit/s, respectively, as evaluated by Temnikov et al. [28]. More recent results for these human sensor channels are similar – about 10^5 bit/s and about 10^7 bits/s [29]. Our calculated capacity values are larger by 4–5 orders of magnitude.

How such a large difference can be explained? First of all, we have calculated the maximum capacity values of a natural lecture, which serve as upper bounds of sound and light sub-channels. This implied that all cells and all frames were equally probable. Practically this is not the case because the sensitivity of ears and eyes are spectrally selective. For example, the human ear is the most sensitive to the sounds in the frequency range from 1500 to 4000 Hz, but the human eye is the most sensitive at the green–yellow light wavelength of 555 nm ($5.4\text{x}10^{14}$ Hz) [18]. Also, the content of a lecture can influence the frame probabilities. In our calculations, we have also neglected the presence of noise in both sound and light sub-channels. Finally, the perceived light and sound information transmission in a nervous system and its processing in the brain are neglected. It is known that a huge information compression takes place there [29]. Evidently, optical and acoustical perception systems of humans are not able to perceive all the physically available information.

Thus, the obtained results for a natural lecture are overestimated. Yet, they can be used as upper limits for corresponding quantities of a video lecture because the information characteristics will be much lower due to the technical limitations of the video camera.

The proposed calculation method of Shannon information characteristics can be used not only to find their maximum values but also, in more general case, to introduce the probability distributions of cells in all frames and also to vary the cell size. The appropriate variations of probability distributions and cell sizes would enable one to meet the empirical conditions [15] of an optimal lecture. Thus, information characteristics of an optimal lecture could be calculated. It should be noted that the mathematical modification of the presented approach is needed in this case.

The method based on the capacity formula (3) of continuous communication channel we have proposed in Section 2 and used to calculate the maximum Shannon information characteristics of a natural lecture is not precise. However, we believe that the method is logical and the made approximations do not dramatically change the results. To prove its practical applicability, further theoretical and experimental studies are necessary.

In principle, the proposed method is quite general. It can be also applied for the calculation of syntactic optical and acoustical information characteristics of any object in the world, e.g., landscapes and streets with people, cosmic objects, etc.

5. CONCLUSIONS

1. The method has been developed to calculate the absolute maximum values of Shannon information characteristics of a natural lecture based on the capacity formula of continuous communication channel and physical considerations taken into account for the first time, to our knowledge. After some modification, it can be practically applied for the optimization of both natural and video lectures because there is some correlation between syntactic and semantic information characteristics.
2. Maximum Shannon information characteristics (entropies of sound and light frames, amounts of total acoustical and optical information, capacities of sound and light sub-channels, total amount of information and total capacity) of a natural lecture perceived by the audience have been calculated. These values are the upper bounds of a video lecture.
3. It has been found that physically the maximum capacity of sound sub-channel $C_{s\ max} = 6.58\times10^9$ bits/s is almost negligible compared to the maximum capacity of light sub-channel $C_{l\ max} = 1.54\times10^{11}$ bits/s. It is the consequence of neglecting meaning in the syntactic information.
4. Capacities of sound and light sub-channels are by 4–5 orders of magnitude larger than the previously estimated capacities of human hearing and vision information channels because of the used approximations and due to neglecting the physiological processes of the information transmission in nervous system and its processing in the brain.
5. Further theoretical and experimental work is needed to develop and prove the proposed method.

REFERENCES

1. Vajapeyam, S. (2014). *Understanding Shannon's Entropy Metric for Information*. eprint arXiv:1405.2061
2. Kolchinsky, A., & Wolpert, D.H. (2018). Semantic Information, Autonomous Agency and Non-equilibrium Statistical Physics. *Interface Focus, 8*, 20180041. http://dx.doi.org/10.1098/rsfs.2018.0041.
3. Carlson, Bruce A. (1986). *Communication Systems. An Introduction to Signals and Noise in Electric Communication.* New York: McGraw-Hill Book Company.
4. MathWorks. (n.d.). *Entropy.* Available at https://se.mathworks.com/help/images/ref/entropy.html
5. MathWorks. (n.d.). *Signal.* Available at https://www.mathworks.com/helps/signal/ref/pentropy.html#mw_da069ed5-a376-4d11-84d3-ca16946eab9
6. Seeling, P. (2010). Scene Change Detection for Uncompressed Video. *Technological Developments in Education and Automation*. doi:https://doi-org.resursi.rtu.lv/10.1007/978-90-481-3656-8_3
7. Xuguang Zhang, X. S. (2019). Crowd Panic State Detection Using Entropy of the Distribution. *Physica A: Statistical Mechanics and its Applications, 525* (7), 935–945. doi: 10.1016/j.physa.2019.04.033
8. Luo, Z. B. (2016). Human Abnormal Behavior Detection Based on RGBD Video's Skeleton Information Entropy. *Lecture Notes in Electrical Engineering.* doi:https://doi-org.resursi.rtu.lv/10.1007/978-3-662-49831-6_74

9. Ferreira, T.A.F.R (2014). *Entropy Based Dynamic Ad Placement Algorithms in Video Advertising*. PhD Thesis, University of Beira Interior, Covilha, Portugal.
10. Nowosad, J. & Stepinski, T.F. (2019). Information Theory as a Consistent Framework for Quantification and Classification of Landscape Patterns. *Landscape Ecol., 34*, 2091–2101.
11. Černekova, Z., Nikou, C., & Pitas, I. (2002). Entropy Metrics used for Video Summarization. *Proceedings of the Spring Conference on Computer Graphics*, Budmarice, Slovakia, 73–82.
12. Liu, J., Wang, Sh., Ma, Wei-Chiu, Shah, M., Hu, R., Dhawan, P. & Urtasan, R. (2020). Conditional Entropy Coding for Efficient Video Compression. *Image and Video Processing*. ECCV, LNC3, 12362, 453–468. https://doi.org/10.1007/978-3-030-58520-4_27.
13. Sun, J., Xu, Zh., Liu, J., & Yeo, Y. (2011). An Objective Visual Security Assesment for Cipher-Images Based on Local Entropy. *Multimed. Tools Appl., 53*, 75–95.
14. Fei, M., Jiang, W., & Mao, W. (2017). Memorable and Rich Video Summarization. *J. Vis. Comun. Image Represent. 42* (C), 207–217.
15. Alksne, L. (2016). How to produce video lectures to engage students and deliver the maximum amount of information. *Proceedings of the International Scientific Conference "Society. Integration. Education",* 503–516. doi: http://dx.doi.org/10.17770/sie2016vol2.1424
16. Wieling, M., & Adriaan Hofman, W.H. A. (2010). The Impact of Online Video Lecture Recordings and Automated Feedback on Student Performance. Computers & Education, 54 (4), :992-–998. doi: 10.1016/j.compedu.2009.10.002
17. Pauliks, R. (2016). *Quality Studies of Video Transmission Services in Packet Networks*. Summary of PhD Thesis. Riga: RTU Publishing House. (in Latvian).
18. Jehonovičs, A. (1984). *Handbook of Physics and Technics*. Rīga: Zvaigzne. (in Latvian).
19. Smith, J. (2007). *Why Can an Opera Singer be Heard over the Much Louder Orchestra?* Available at https://www.scientificamerican.com/article/expert-opera-singer/
20. Horowitz, S. (2013). *The Universal Sense: How Hearing Shapes the Mind.* USA: Bloomsbury.
21. Nave, C.R. (2016). *Hyperphysics.* Atlanta: Georgia State University.
22. Errede, S. (2002–2017). The Human Ear-Hearing, Sound Intensity and Loudness Levels. *UIUC Physics 406 Acoustical Physics of Music*,1–33.
23. Benjamin, A.T., & Quinn, J.J. (2003). *The Proofs that Really Count. The Art of Combinatorial Proof*. The Dolciani Mathematical Expositions 27, The Mathematical Association of America, ISBN 978-0-88385-333-7.
24. Gaisler, W.S, & Banks, M.S. (2010). Visual perfomance. In M. Bass, *Handbook of Optics, vol. III. Vision and Vision Optics* (pp. 2.1–2.51). New York: McGraw Hill Companies, Inc.
25. Luizov, A.V. (1989). Colour and Light. Leningrad: Energoatomizdat. (in Russian).
26. Richards, A. (2011). *Alien Vision. Exploring the Electromagnetic Spectrum with Imaging Technology.* (2nd ed.). Bellingham, Washington: SPIE Press.
27. Werner, J.S., Schefrin, B.E., & Bradley, A. (2010). Optics and Vision of the Aging Eye. In M.Bass, *Handbook of Optics, vol. III. Vision and Vision Optics* (pp.14.1–14.38). New York: McGraw Hill Companies, Inc.
28. Temnikov, F.E., Afonin, V.A., & Dmitriev, V.I. (1971). *Theoretical Foundations of Information Technics.* Moscow: Energija. (in Russian).
29. Markowsky, G. (2017). *Information Theory*. Encyclopaedia Britannica. Available at https://www.britannica.com/science/information-theory