Redundancy versus Video and Audio Human Perception

MICHAEL BANK
Holon Institute of Technology (HIT), Israel
michaelbank@bezeqint.net

Abstract— A proposal is put forward for an possible explanation of human perception algorithm. The main conclusion of different kinds of art perception examination is necessity of big part of signal redundancy. The article illustrates the meaning of redundancy versus human perception. It turned out, that there are common methods of redundancy creation in different kinds of art like painting, architecture, music and speech.

Key-Words: - Painting, music, speech, perception, redundancy, Autism

1.Introduction. Practical meaning of perception algorithm study.

Since the invention of the radio, there has been an exponential increase of transmitted announcements and entertainment programs. Due to the limited bandwidth in communication channels, transmitted signals must be compressed. The most stubborn problems are found in mobile communication systems where in order to prevent channel problems, the size of the signal (bit number) is increased, for example, in Multipath Propagation and Doppler Effect systems. Typical examples include Error Correction Code (ECC), pilot signals, headers in pocketing signals, etc. It is shown the common peculiarities of human processing of video and audio information. On the other hand, due to the limited bandwidth in communication channels, transmitted signals must be compressed. This begs the question – what should be transmitted in these types of signals, if a large part of them is removed by the system controlling the sensor-brain?

A human being cannot interpret or remember all signals that are received by the eyes and ears. If the brain was capable of absorbing every bit of information the brain would reach maximum capacity soon after birth and would not be able to absorb additional information. There are effective compression techniques for removing a large proportion of received signals.

This begs the question – why transmit these types of signals if they are removed by the system controlling the sensor-brain?

Since we do not know the manner in which the brain receives impressions, knowledge of algorithm perception enables us to invent new methods for decreasing noise level via modification of bits into bits that do not contain noise and that do not weaken the level of perception without deterioration in the level of perception.

A study of algorithm perception will enable the development of new methods for video and sound signal transmission quality. All known methods are based on the measurement of signal variation. Presently, this is not possible due to deliberate signal variation during the compression process.

However, knowledge of algorithm perception enables one to do things such as transferring certain features of a work-of-art onto a computer.

2. Information and redundancy of video and audio signals

2.1 Statistical redundancy.

An announcement source signal contains redundancy. We can elaborate by declaring that the announcement signal consists of information and redundancy [1-5]. We can elaborate by declaring divide redundancy into two components: statistical and psychological. Information can also be divided into two components: coloring and meaning (see Fig. 1).

If there is a correlation between signal (A) samples, statistical redundancy exists, which is possible to

 Manuscript Received Jan.30, 2008; Revised March 26, 2008

evaluate by (1), where K is the number of different possible samples, and H is signal entropy.

$$\chi = \frac{\log K - H(A)}{\log K} \cdot 100\% \qquad (1)$$

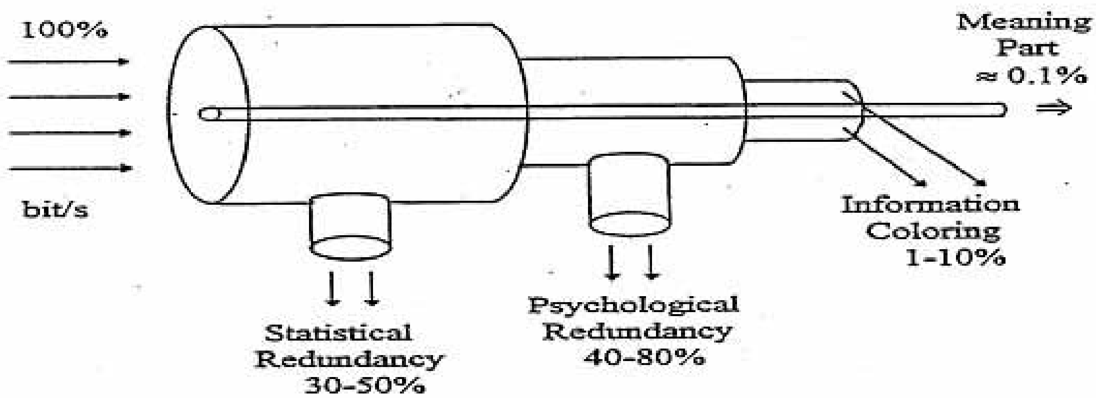


Figure 1. Signal structure

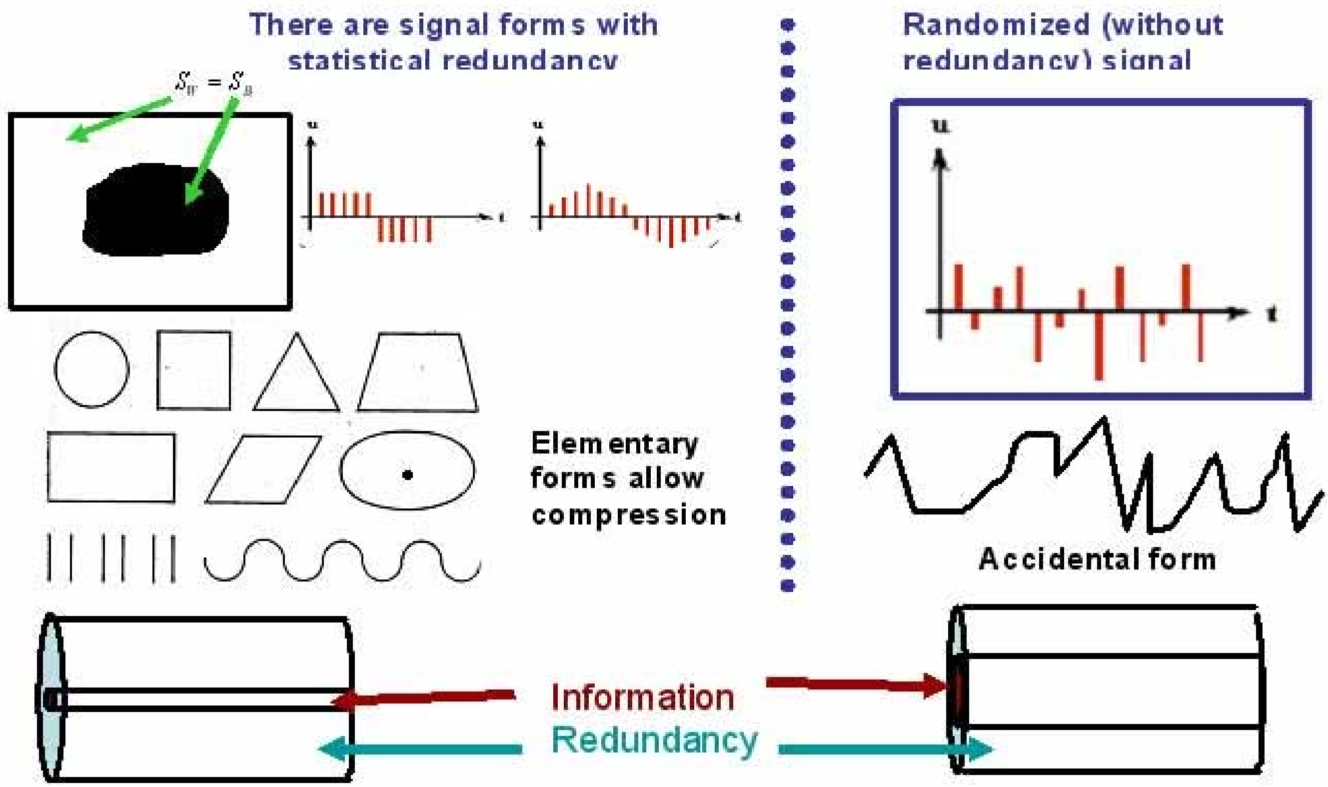


Figure 2. Signals with different statistical redundancy.

White noise consists of random signal without redundancy, since there is no correlation between its samples. Large statistical redundancy enables you to predict the following signal elements. Examples of signals with and without statistical redundancy are illustrated in Fig. 2.

Deleting statistical redundancy from signals does not prevent full signal recovery. Deletion and recovery of statistical redundancy is termed lossless method. A well-known example of statistical redundancy reduction is the Morse code.

The following example illustrates the meaning of redundancy. Figure 3a consists of a clear and animated image that has large redundancy. If we rearrange the location of the pixels we will arrive at a picture (Figure 3b) with the same number of bits but with very small redundancy and which appears jumbled and unclear. This brings us to conclude that redundancy is proportional to the clarity of an image.

Figure 3a - Original animated and clear picture with large redundancy.

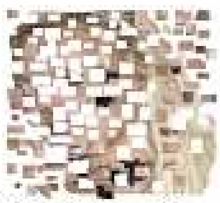


Figure 3b - Jumbled picture with small redundancy.

.

2.2 Video signal physiological redundancy examples

Physiological redundancy consists of part of the signal that is not required for perception. It is a proven fact that the system controlling the brain-sensor removes a large proportion of information that enters the eyes and ears. Removing physiological redundancy in the encoder does not mean there will be full signal recovery in the decoder. However, this type of signal variation is not usually noticeable by an audience.

We will now examine several examples of physiological redundancy found in video and audio signals.

A television camera transmits information on each pixel from line to line [6,7]. The order in which our eyes view this process is: all contour lines; then, the curved loops; and finally the central points. There is considerable “shaking” along contour lines. Subsequently, the eye transmits less quantity of information from positions without clear lines (see Fig. 4).

In other words, a decrease in bit number corresponds to a section of the picture with no clear lines

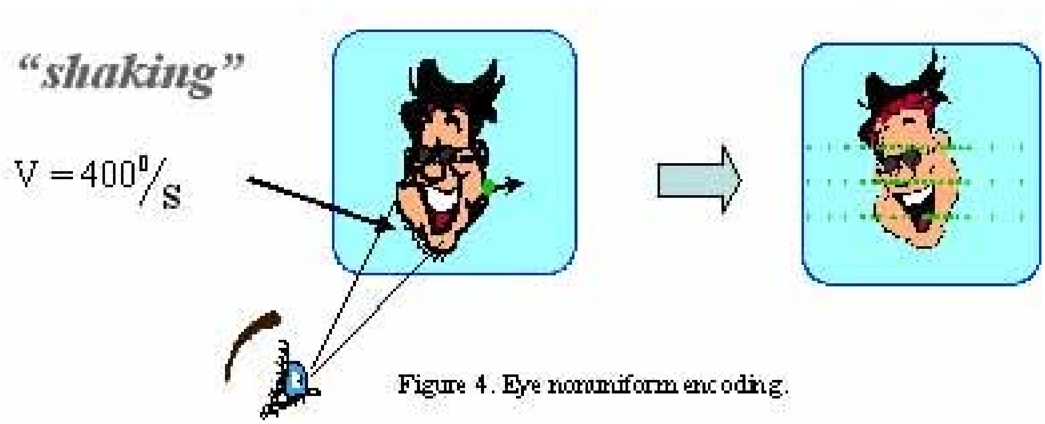


Figure 4. Eye nonuniform encoding.

Another classic example of physiological redundancy is the observation of moving objects. Only 15 pictures per sec are needed to achieve continuous movement in a moving object. Where the velocity is great, only discrete moments are needed by the system controlling the brain-sensor (See Fig.5).

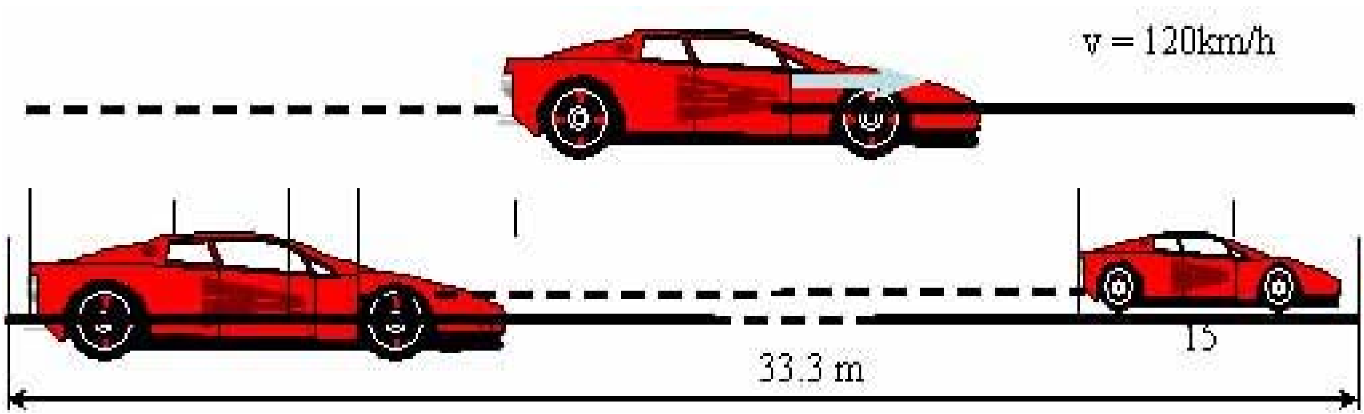
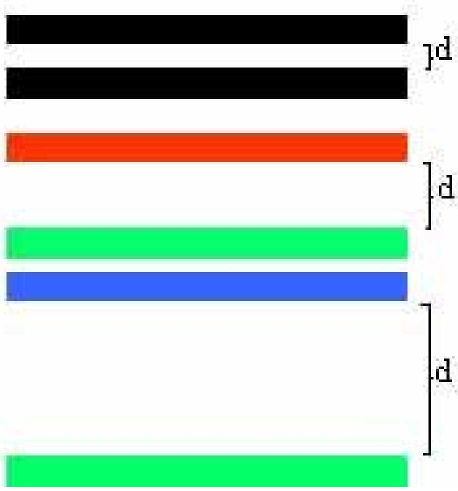


Figure 5. Moving perception.

Another example of physiological redundancy is color resolution where resolution is smaller than black-white resolution. Through black-white resolution one can view finer details (3-4 pixels). For example, where there are two separate black lines the distance between the lines need only be 1 mm. However, for color lines the distance will be greater (See Fig 6.). Within the entire range of existing colors, human beings can differentiate only 180 different colors.



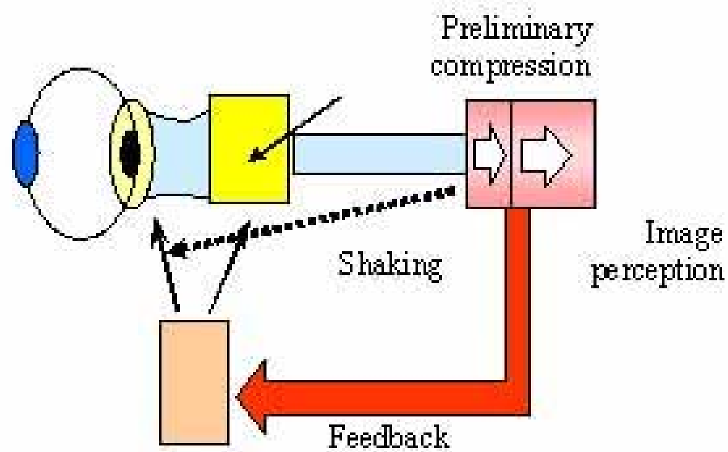


Figure 6 Lines resolution on the same distance.

Figure 7 Model of seeing control system

2.3. Vision perception system models.

In the case of video signals, compression is more effective when searching for objects which require more detailed transmission (including contours, lines, central points, and periodic forms).
Communication between the brain and the eye's pupil enables the brain to remove statistical redundancy and to gain control upon the movement of the pupil for achieving image perception through a minimum number of bits (see Fig. 7).
The eye can receive a great amount of information (more than 1 Gbit/s). However, only a small part of this information is received by the brain (approx. 100 bit/s).

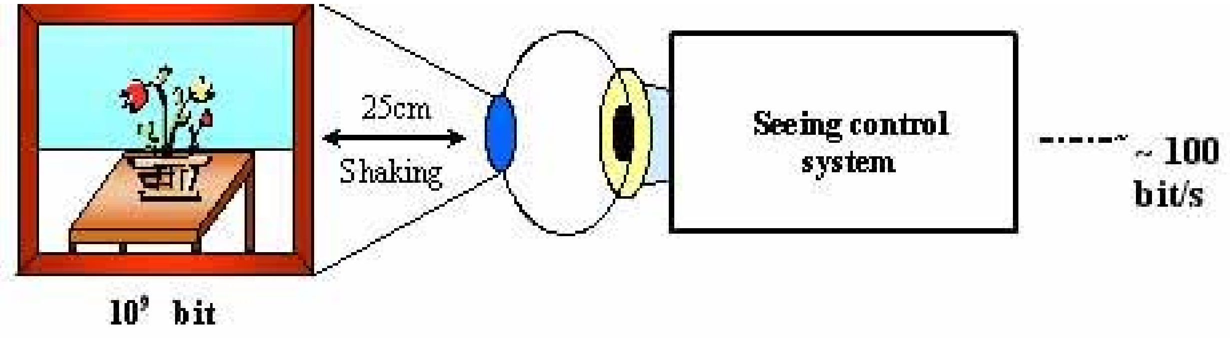
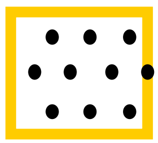


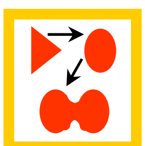
Figure 8. Compression intensity on human being eyesight system

Why only 100bit/s ? Numerous psychological experiments indicate:

During a period of 0.7 seconds the brain calculates the details of an image at the rate of 70 bits/s.



We can perceive different parts in a painting within a given time frame of less than 0.8 s. The brain calculates this at the rate of 80 bits/s



Maximal reading speed is 30 – 40 bit/s

The conclusion here is that the brain receives impressions at the rate of less than 100 bits/s

2.4 Audio signal psychological redundancy examples.

Sounds cannot be heard 100 ms above volume peak and 5 ms below volume peak. Furthermore, during this period of time no signal may be transmitted.

One cannot hear small spectral components, whereas maximal components can be heard because of a significantly greater difference in frequency(see Fig. 9).

This is termed Masking Effect, which is used in most audio signal compression methods [8-11].

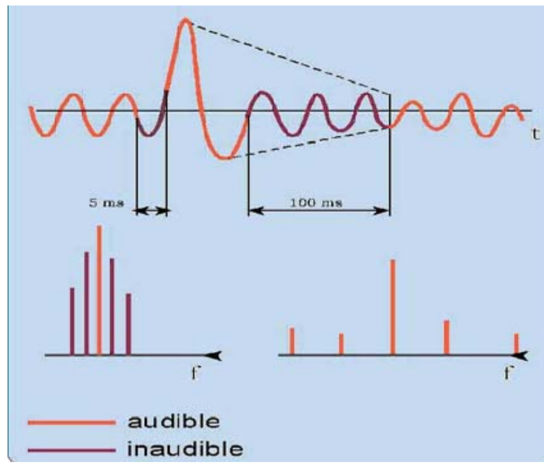


Figure 9 Psychological redundancy in time domain and in frequency domain.

]The Masking Effect has been researched extensively. Critical frequency band tables divide bandwidths into critical frequency bands where the Masking Effect is used (see Fig. 10). Following spectral transformation of real sound signals, we can transmit only one or two maximal components of each critical frequency band.

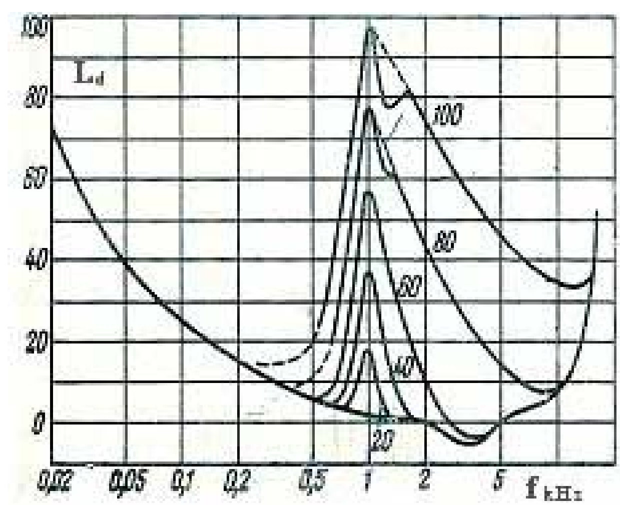


Figure 10 Masking Effect experimental data. The case of one spectral component (1000 Hz)

2.5 Model of Human Hearing Perception

Let us assume audio signal consists of short stationary fragments, as illustrated in Fig. 11. These fragments vary after time t1, t2, etc. If we take other fragments according to constant times T1, T2, …, as in all of today's audio compression systems, we will obtain complicated spectrums. In this case, transfer takes place from one stationary signal to the other.

If we take stationary fragments (times t1, t2, …) we will obtain simple spectrums. Subsequently, the information quantity of signal in spectral form depends on choosing the method of time interval.

Maybe this feedback from brain to hearing processor is explanation of strong rhythm influence.

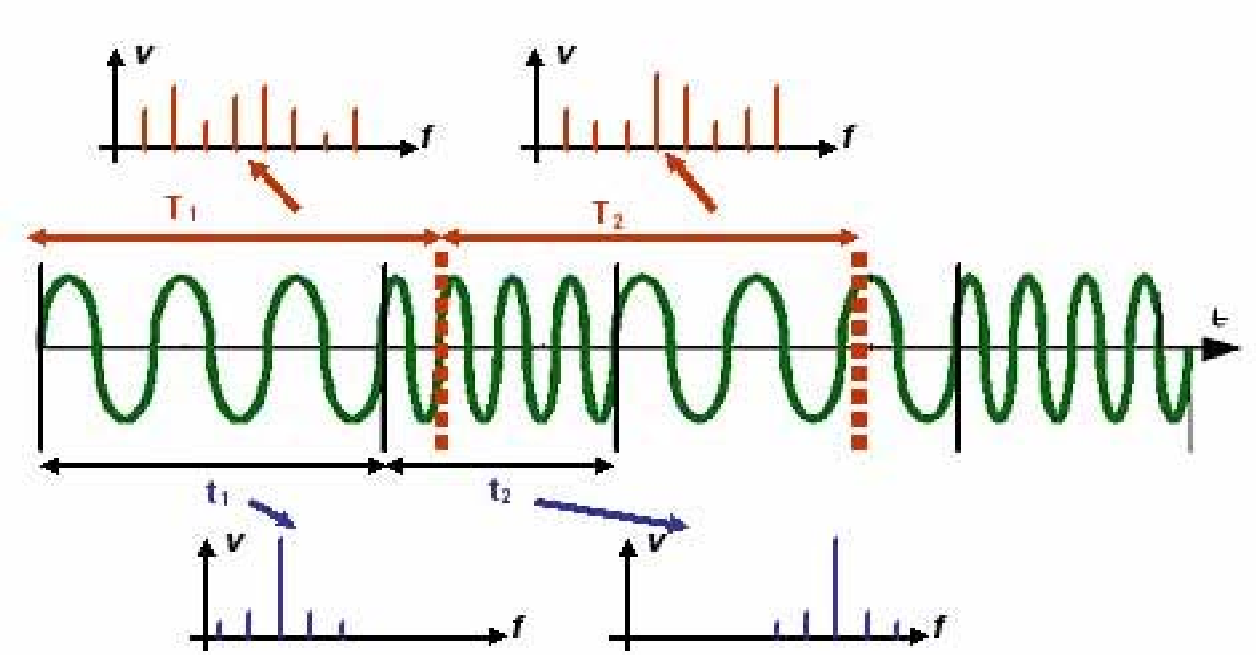


Figure 11 Time intervals choosing.

The following are three important points worth considering:

- The ear membrane consists of 23000 resonators and operates similar to a spectral analyzer.
- The brain receives a small amount of spectral components.
- There can be no music or even speech without rhythm. In addition, rhythm is a quality that we physically sense and is not only something we hear.

We propose a Hearing System model (see Fig. 12), which operates within frequency and time domains, simultaneously. The system divides signals into fragments corresponding to stationary signals. The system then conducts spectral transformation and finally leaves only the maximal components in each critical frequency band.

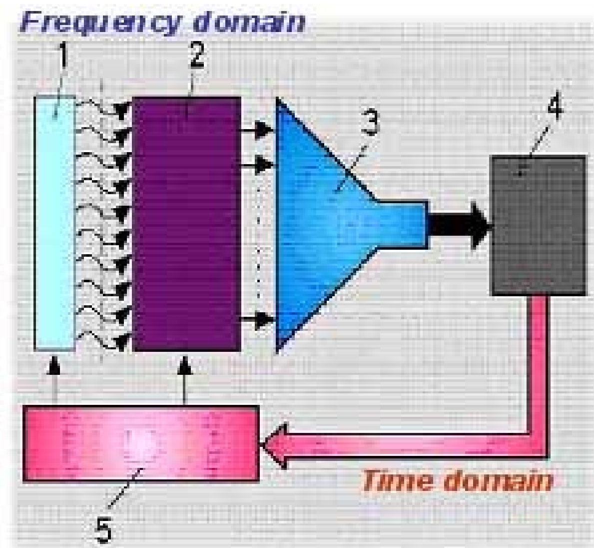


Figure 12 Hearing system model
1 - resonators, 2 – Encoders, 3 – Channel between Ear and Brain, 4 – Central Processing Unit, 5 – Spacing feedback.

3. The first hypothesis

The more a work-of-art, whether it be music, painting, or architecture, corresponds to human vision and to system models of hearing perception, the more enhanced will be the work-of-art.
The brain's volume (cell) is limited in receiving information (see Fig. 13).
The receptor-brain system removes redundancy.
The tendency of the receptor-brain system is to compress received signals even if the data is of a greater size than the cell volume of the brain. This dissonance gives us a sense of uneasiness.
For greater reception of impressions and agreeable perceptions, we must achieve maximal statistical redundancy of input signals.
For this purpose, the fine arts implement predictable lines, elementary forms, recurrent (rhythmic) figures, predictable colors, and repetition of some of the fragments. Music implements clear rhythm together with frequencies which correspond to the critical frequency band center.
Our memory system operates according to patterns or samples, as follows: closed loop with emphases center; separate form; construction part; melody; and rhythmical picture. If we want to store image information then fewer bits are required as opposed to memorizing separate bits. In information theory, this compression method is termed “Alphabet Extension” or “Using Code Book”.
People usually wait for patterns to be perfected through repetition since this helps in better remembering images. A small variation in repetition - slight difference in the eyes, nonsymmetrical face, nonsymmetrical build, slight variation in rhythm or volume repeated melodies or phrases - can produce the strongest impressions.
We can find confirmation for these theories in famous works of art. Artists, sculptors, architects, and composers have emphasized, by intuitive perception, the peculiarities above. Numerous examples are listed below.

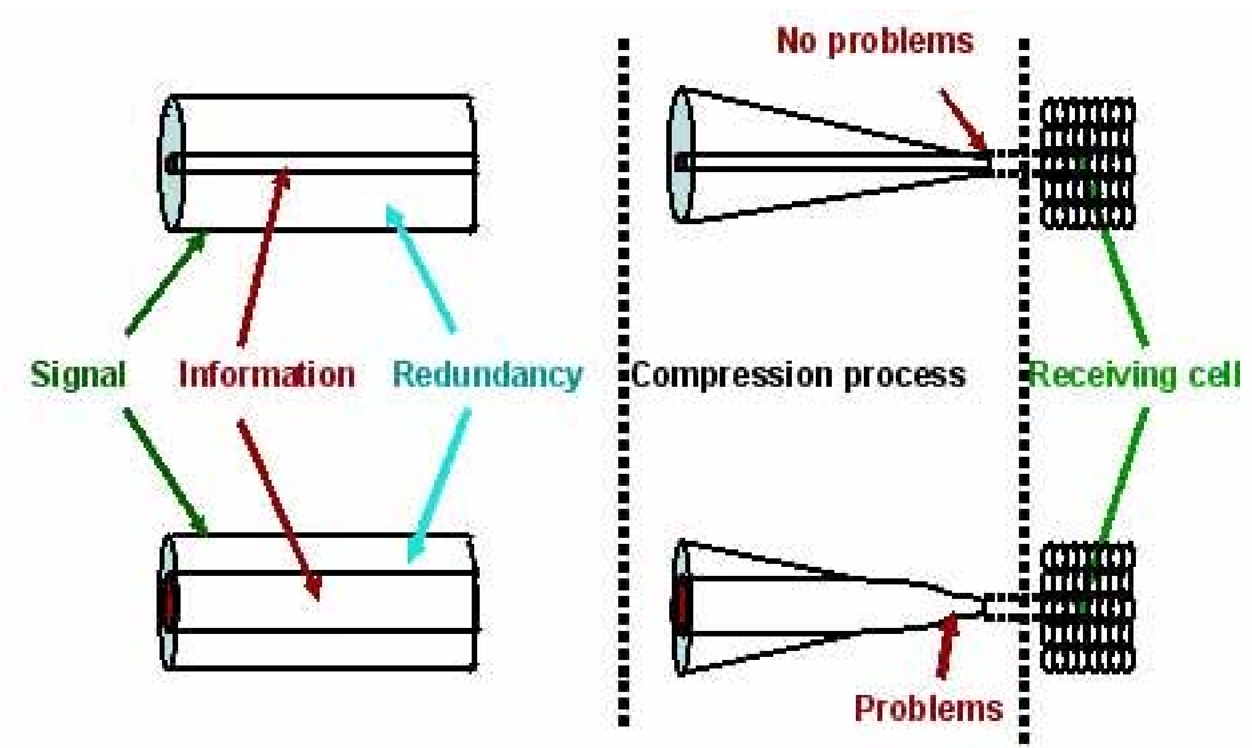


Figure 13 Big redundancy advantage.

4. Increasing redundancy in visual signal

To decrease the number of bits for transmitting the visual signal and for increasing statistical redundancy, it is necessary to implement predictable lines and figures (that have large autocorrelation).

It is a known fact that the ratio of size dimensions corresponds to Fibonacci numbers,

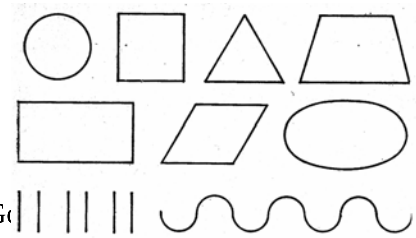
$$F(n) = \begin{cases} o & n = 0 \\ 1 & n = 1 \\ F(n-1) + F(n-2) & n > 1 \end{cases}$$

or to additional serial,

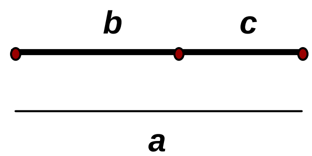
$$1 \quad 2 \quad 3 \quad 5 \quad 8 \quad 13 \quad .... \qquad \frac{a_n}{a_{n-1}} \approx 1.618$$

Throughout history, the ratio for length to width of rectangles of 1.61803 39887 49894 84820 has been considered the most pleasing to the eye. This ratio was named the golden ratio by the Greeks. In the world of mathematics, the numeric value is called "phi", named for the Greek sculptor Phidias. The space between the columns form golden rectangles.

Great Pyramid of Giza

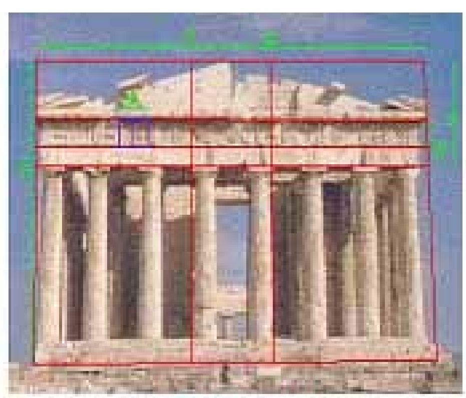


or to G

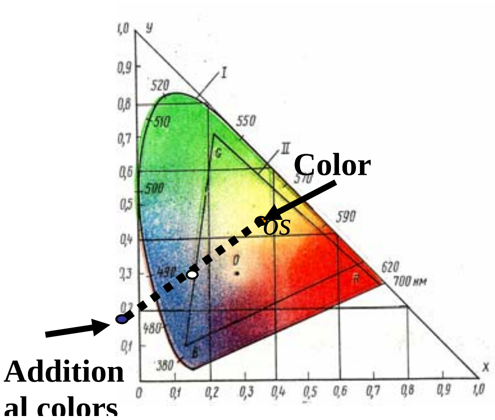


$$\frac{b}{a} = \frac{c}{b} = k \approx 1.618 \qquad b = \sqrt{a \cdot c}$$

There are golden rectangles throughout this structure which is found in Athens, Greece. It is known that the ratio taken for half of the perimeter size and the pyramid height equals π.
Use of the Gold section increases redundancy, due to high predictability of neighboring lines or form.



Parthenon



Conjugate (coupled, additional) colors is another source of redundancy Implementation of conjugate colors enable the use of fewer number of bits, since these colors correspond to OS^1 = OS, where point 0 is white color.

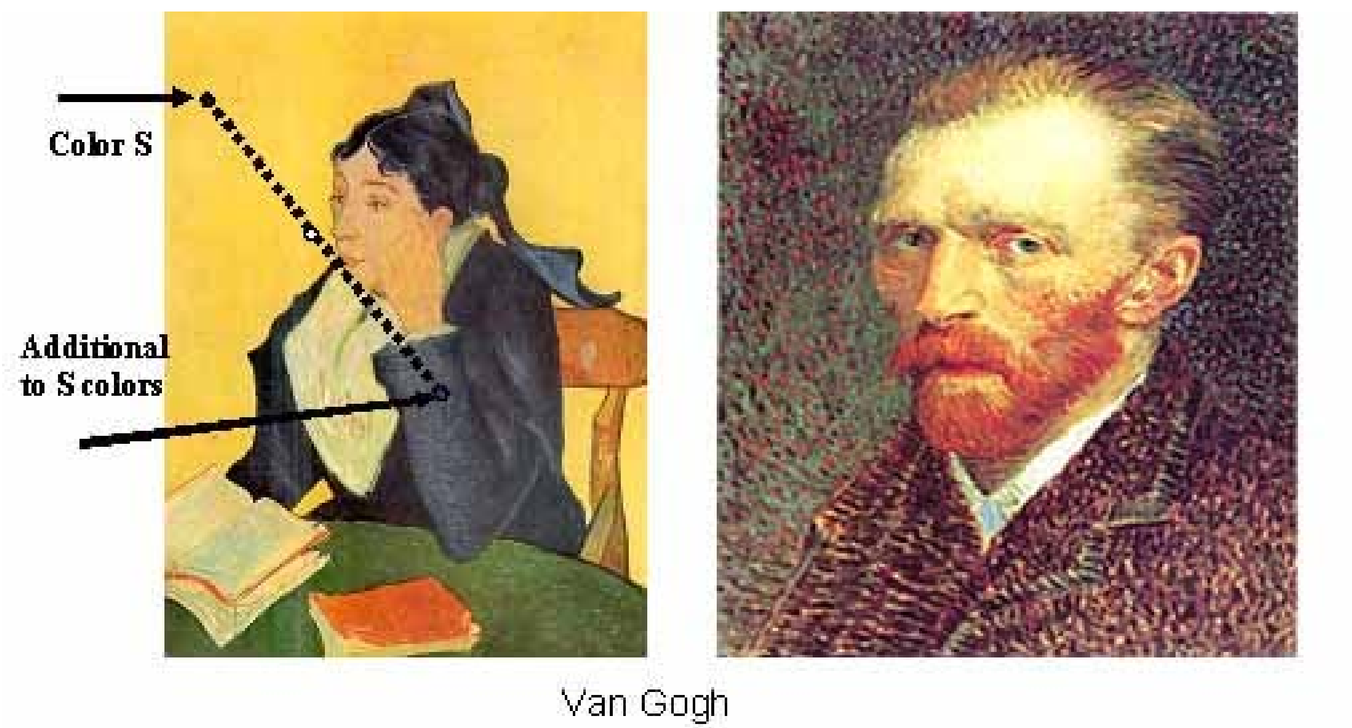
Figure 14. Colors locus

5. The Greatest Artists “know” this theory.

The portrait seen here and the following three portraits confirm our proposed hypothesis.

The accented part includes a small difference in the eyes as in a closed loop.

The remainder of the portrait consists of repeated fragments with large statistical redundancy.



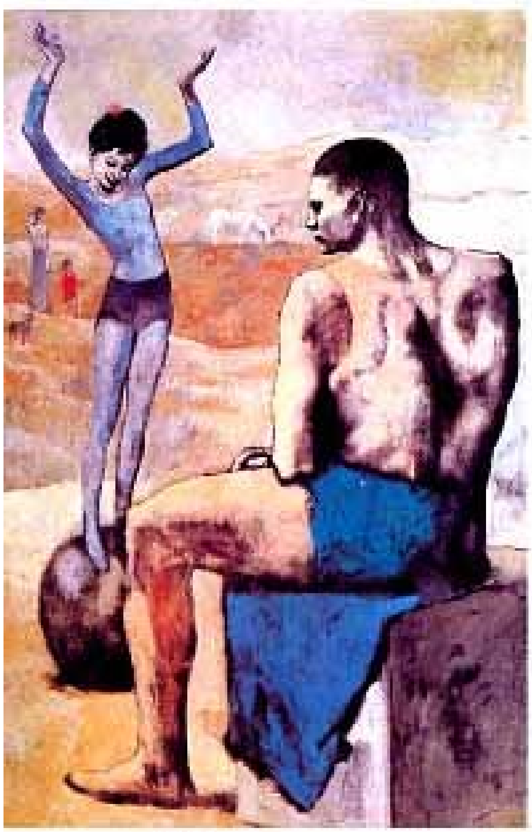
Renoir

This is a portrait of Natalie Pushkin. She was considered to be the first beauty in Russia during the first half of the nineteenth century. It was said that she was a little cross-eyed and this feature is said to have made her more attractive.

Brulov

Modigliani

Henri Matisse



Picasso.

Professional painters use contour lines to influence and sway people's perception of a portrait. These lines also contain large redundancy.

Each women "knows" this theory too.
Considering that the purpose of putting on makeup is to attract attention, the picture shown here and in the following slide displays accented closed loop lines with the emphases again on the center.

Inga, expert beautician

6. In architecture one must achieve maximum redundancy to thrill the observer.

Florence, Sant Miniato, XIII sent.

St. Isaak`s Cathedral, St Petersburg.

St Petersburg University.

In this picture and in the following three pictures you can see predictable lines (including Golden Section), elementary forms, clear contours and recurrent (rhythmic) figures and lines.

Repeated forms and clear rhythms (which indicate large redundancy) generate the strongest impression.

An early masterpiece of architecture and nature is the Pokrov in Neril.

7. The Greatest Composers "know" this theory too

The strings sound overtones. Overtone is sound with of double frequency. For example f 2f 4f 8f …

In compliance with this, all sounds (notes) are divided by octaves. Each melody can be performed in each octave. Inside of each octave we have m = 12 tones. Out of 12 tones, 8 tones are used which correspond to tonality.

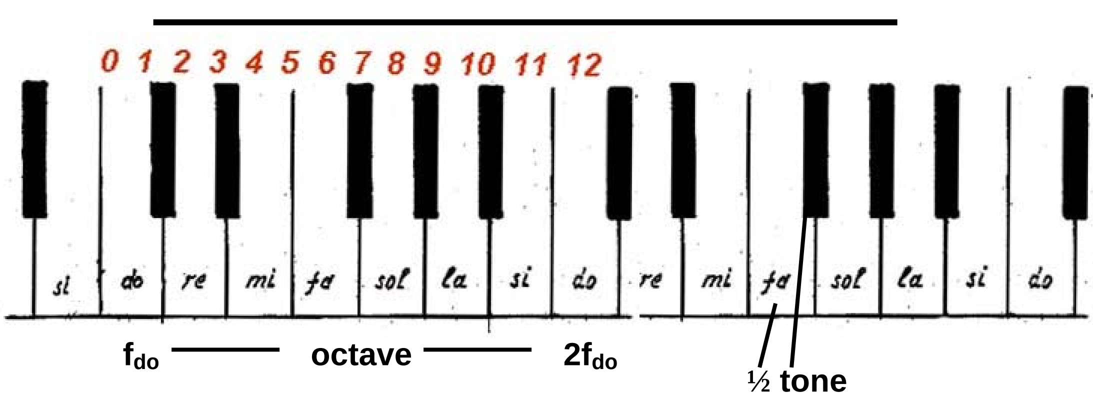
From year 1700:

$$\frac{1}{2} tone = \frac{f_1}{f_0} = \sqrt[12]{2}$$

$$1 tone = \frac{f_2}{f_0} = 2\sqrt[12]{2} \quad \log \frac{f_2}{f_0} = 0,167$$

$$Octave = \frac{f_{11}}{f_0} = 2$$

Using constant rates between tones and dividing by octaves increases redundancy significantly.



Accords (simultaneously sounding tones) result in a greater increase in redundancy, since a rhythmical sound can be heard only through a limited combination of notes. All accords can be divided by Consonances (C) and Dissonances (D). C consists of a pleasant sound. D's appearance in real musical composition sounds like something impossible and unpleasant. The explanation below paragraph provides a proposal for this phenomenon.

N	Name	f_1/f_2	$\frac{f_2-f_1}{\sqrt{f_1\cdot f_2}}$	$\frac{2f_1-f_2}{\sqrt{2f_1\cdot f_2}}$	C or D
1	S second	16/15	0.065	0.64	D
2	L second	9/8	0.12	0.58	D
3	S third	6/5	0.18	0.52	C
4	L third	5/4	0.22	0.47	C
5	Quart	4/3	0.29	0.41	C
6	Quint	3/2	0.41	0.29	C
7	Lsixth	5/3	0.47	0.22	C
8	S sixth	8/5	0.52	0.18	C
9	S seventh	16/9	0.58	0.02	D
10	L seventh	15/8	0.64	0.065	D
11	Octave	2/1	0.707	0	-

$$D \rightarrow \left\{ \frac{f_2-f_1}{\sqrt{f_1\cdot f_2}} < 0.16 \quad \text{or} \quad \frac{2f_1-f_2}{\sqrt{2f_1\cdot f_2}} < 0.16 \right\}$$

In the case of D, where there are two notes, or, one note together with a second note overtone fall into one critical frequency band. The "Masking effect" must decrease the smaller component in one band, but, as in the case of D, both components have the same volume. The critical frequency bandwidth equals approximately 16% of central frequency. The proof of this explanation is illustrated in this table.

A musical composition consists of melodies. Melody has a large autocorrelation. This allows us to predict subsequent tones in the melody. Thus, a melody signal has large redundancy. Similar to the pattern of a picture, we remember a melody by repetition of the same melody. In other words, the melody is played not only once, but many times over. Both the composer and performer slightly change the melody during repetition, without a variation in "content". The following parameters can be changed: volume, use of instruments, micro pauses inputting, etc. These parameters increase redundancy and help to create a greater impression.

There are many musical examples of prominent artists who have made successful use of melody repetition. For example, Revel's Bolero or in the first section of Shostakovich's Seventh (Leningrad) Symphony.

Rhythm holds great importance in the construction of a musical composition. There can be no music without rhythm. Rhythm helps us to perceive audio signals like in stationary fragment sequences with large redundancy. Subsequently, poetry, which is defined to a large extent by rhythm, has greater influence on people than simple text.

As a matter of fact people look for rhythms in all kinds of sounds, for example, in the swishing sound of a car's windscreen wipers, or in communal prayer, or when listening to the tweets of birds.

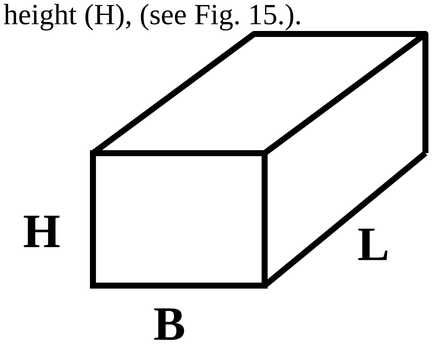
Here we can make rough estimation of human perception audio signal compression rate. We can hear K = 23000 different frequencies. Maximal entropy equals $\log_2 23000 \approx 14.5$

Notes number is 240.In one tonality it will be 120.

Notes in middle octave have maximum probability. Notes correspond to basic tone (tonic) have the highest probability.

The entropy of notes distribution is approximately 0.1. If we take into consideration differential method for transmitting seriate melodies we will get $H(A) \approx 0.02$ or $\chi \approx 99.8\%$. So compression rate is ≈ 500.

It is interesting to note that there is another item which is common to painting and music. To achieve good acoustic conditions a room is required with Golden Section between length (L), width (B) and height (H), (see Fig. 15.).



$$L/B = B/H, \quad L = B + H$$

$$L = 1{,}62\sqrt[3]{V} \quad B = \sqrt[3]{V}$$

$$H = 0{,}62\sqrt[3]{V} \quad (V = LBH)$$

Figure 15. Preferable sides ratio for good acoustic

8. How to give an effective speech.

Wikipedia: ***Oratory is the art of public speaking. In ancient Greece and Rome, oratory was studied as a component of rhetoric (that is, composition and delivery of speeches), and was an important skill in public and private life. Aristotle and Quintilian discussed oratory, and the subject, with definitive rules and models, was emphasised as a part of a "complete education" during the Middle Ages and Renaissance.***

Below are special rules for putting together a speech, which are original and unexpected in their nature.

- After attending a lecture, the average person will be remember about 10% of what was talked about. Therefore, it is of paramount importance that this 10% includes the main idea you wish to get across to your audience.
- Do not begin your speech blatantly disproving present theories, which are generally accepted. Rather, mention that present theories hold valuable information, and then mention your innovative method.
- In the first quarter of the lecture prove your ideas by expounding them. Next, give examples, analogies, and mention other peoples' opinions in order to bring increased clarity to these new ideas.
- Try to be as convincing as possible when giving a lecture, even to the point where the listener asks himself "yes, this makes sense. Why didn't I see this before".
- When a listener finds it easy to absorb and understand what is being said, he is more likely to go along and to agree to what you are saying and putting forward in your speech.
- Good organization is vital and it is best to plan and write up what you intend to say in your speech. Nevertheless, do not read out the speech word for word. Rather, if you have carried out your planning correctly, the words in your speech should flow smoothly.
- Try not to include superfluous words. Do not make repetition of words and phrases, such as: "well", "basically", "the truth of the matter is", etc. Come straight to the point, thereby keeping the wording of the speech to a necessary minimum.
- It is possible to increase volume a little and stress rhythm if your argument is weak. If you find that your argument is rather weak and lacking in depth, then speaking up will get your point across more effectively.
- If you have a complex and difficult point you want to get across to your audience, how about telling a joke or anecdote straight after.
- Look at the faces of the members of the audience. Do not let your eyes wander or gaze on the walls or on what's occurring outside the window.
- Conclude your lecture by repeating the main ideas.
- The new information presented in your speech should represent 10% of what you say. The rest is redundancy. Therefore, it is worth your while to stress and to further expound upon the main 10% of your speech.
- Try to think of something influential to say for your closing sentence. The audience pays particular attention to the final words that are uttered by the orator.

The above points are not new. In fact, they were recommended long ago. They are a testament to the following common rule: "for success in public speaking one must use redundancy correctly"

9. Autism problems, the second hypothesis

Wikipedia: Autism is a brain development disorder that impairs social interaction and communication, and causes restricted and repetitive behavior, all starting before a child is three years old. This set of signs distinguishes autism from milder autism spectrum disorders (ASD) such as Asperger syndrome. In the second and third years, autistic children have less frequent and less diverse babbling, consonants, words, and word combinations; their gestures are less often integrated with words. Autistic children are less likely to make requests or share experiences, and are more likely to simply repeat others' words or reverse pronouns Joint attention seems to be necessary for functional speech, and deficits in joint attention seem to distinguish infants with autism spectrum disorders (ASD) for example, they may look at a pointing hand instead of the pointed-at object, and they consistently fail to point to "comment" about or "share" an experience at age-appropriate times. Autistic children may have difficulty with imaginative play and with developing symbols into language.

It may be that children diagnosed with autism have difficulties deleting redundancy. Habitual speech for them sounds dissonant. They have difficulties verbalizing their thoughts and often their speech is very simple and lacking depth [12 - 13] **Perhaps we must communicate with these people using minimum redundancy. Instead of detailed and image-bearing descriptions, one can use simple telegraphed text or poem style text. This is especially important in the teaching process.**

10. Conclusion

The information, experimental results, assumptions, and hypotheses within this piece of writing bring us to a paradoxical conclusion:

To increase the influence of a newly released work of art and receive maximal impressions, we must decrease information quantity in signals between the object and observer or listener.

The greater part of the signal is intended for redundancy. To this end:

- We must implement lines, figures, word patterns, and combinations of notes with large autocorrelation;
- Rhythm, which is an essential component in music and speech, is also a desirable quality in paintings;
- Rhythm uses separate patterns, clear contours, and melodies etc, which include finer bits of information.
- It is possible to increase impressions with the help of pattern/sound repetition, especially where the differences between each repetition are minimal.

There is an additional conclusion which may not be considered scientific.

It is conceivable that millions of years ago, as a result of mutation, a distinctive class of monkey entered the universe that was able to analyze received information in a more subtle manner. This monkey was not only able to receive danger signals or to experience various emotions, including "love", but also had the ability for deleting redundancy by its brain system. This culminated in receiving information in a manner which made more efficient use of stored information, and in the enhanced ability of comparison between new and stored information.

Perhaps this mutation contributed to a great extent in the transformation of the monkey into a human being.

Reference:

1. Shannon, C.E. The mathematical theory of communication. Bell Syst. Tech. J., Vol.27, pp.379-423 and pp. 6230656, 1948.
2. R.E. Blahut, "Principles and Practice of Information Theory", Addvison-Wealey, 1987.
3. R.W. Hamming, "Coding and Information Theory", Prentice-Hall, 1980.
4. Jan C. A. van der Lubbe, "Information Theory", Canbridge Univer. Press, 2002.
5. M. Haridim, M. Bank, V. Vulfin . Investigation of the impact of pilot signals on the performance of OFDMA systems. Int. WSEAS J. of Applied Mathematics and Informatics. 1(2). 2007 (84 – 91).
6. Townsend. B. "PAL Color Television" Cambridge University Press, 1970. .
7. Watkinson. J. "The Art of Digital Video" Focal Press, 2000.
8. Zwicker. E. …"Psychoacoustics, Facts and Models", Springer, Verlag, 1990.
9. Bank. M. "Bearbeitung der schallinformation im menschlichen gehorsystem und in technischen anlagen" Rundfunktechnischen Mittelungen, N2, 1992 (pp. 53-65)
10. M. Bosi, R. E. Goldberg, "Introduction to Digital Audio Coding and Standard", Kluwer Acad. Publ. 2003.
11. M. Bank, U. Mahlab "Hearing system model and sound quality estimation" WSEAS Transactions on Acoustics and Music, 1(1), 2004 (34-44) See
12. M. Bank. Video and Audio Compressions and Human Perception Mechanism. Plenary lecture on 2th WSEAS International Conference on COMMUNICATIONS
13. Human perception seminar in www.OFDMA-MANFRED.com