A Sensory Substitution mapping model based on human brain's pattern recognition ability incorporating human visual system parameters fused with multimodal feedback

Abstract :

The human brain acts as a pattern analyzer at the decision making level and that can be leveraged, in case of the sensory substitution problem. The question that we put forward - Is it actually possible to see with our ears or with our skin , might seem bizarre to us at the outset. But after researching for some time on the topic, to find solutions as well as answers to these questions, we could come up with a probable solution. If we tackle this problem as a union of sets related to neuroscience, signal processing, bio-inspired algorithms or systems like human visual system and auditory system while fusing these together, we might end up with a solution for the people deficient in one of the senses. In our case, we focus on the visually impaired people, and give them the experience of visual perception using haptically encoded data or auditory data. Research has shown that the human brain is evolved to understand and decode pattern, be it visual perception or auditory perception. We approach the problem from signal processing perspective and bio-inspired modelling of the human vision and auditory systems. So to extract and classify the most important features from a given visual data feed, we need to study and emulate the human visual system's bio-inspired model. Once we can make an algorithm for emulating the visual perception model, we can derive a weight to express various parameters or factors used by human brain for decision making for visual perception. We can then realize these parameter(s) as the appropriate auditory feedback parameter(s) to be a cross-modal or associative function for the end-user to experience the visual data as an appropriately mapped audio signal.

<u>Overview of tasks in order :</u>

1. Take the Raw Image from any generic image acquisition device (imitating the human eye layer of Human visual system)
2. Change color space (YCbCr, RGB, etc) according to type of external environment & person or target information to be transferred
3. Convert image to frequency space (dct or dwt transform) for better analysis of image features
4. Research and come-up with a model to find important parameters from the transformed image in frequency domain according to a bio-inspired Human Visual System (HVS) model emulating the human perception
5. Extract those HVS parameters from the given image and convert it to a parameter/weight for the audio data feed or haptic data feed [maybe a multimodal feedback system]
6. Convert that HVS modelled information into haptically and auditory encoded data in a cross-modal mapping [Amplitude, frequency, binaural mode maybe]
7. Make use various encoding methods and patterns for improving sensitivity, accuracy of model, while encoding the image signal into auditory space

Appendix :

Base papers:

[1] Novich, S.D. & Eagleman, D.M. Exp Brain Res (2015) 233: 2777. Using space and time to encode vibrotactile information: toward an estimate of the skin's achievable throughput
https://doi.org/10.1007/s00221-015-4346-1
https://link.springer.com/article/10.1007/s00221-015-4346-1

[2] David Daniel Cox, Thomas Dean, Neural Networks and Neuroscience-Inspired Computer Vision, Current Biology, Volume 24, Issue 18,2014,Pages R921-R929,ISSN 0960-9822,
https://doi.org/10.1016/j.cub.2014.08.026
http://www.sciencedirect.com/science/article/pii/S0960982214010392

[3] David Brown, Tom Macpherson, Jamie Wardô School of Psychology, University of Sussex, Falmer, Brighton BN1 9QH, UK; Seeing with sound? Exploring different characteristics of a visual-to-auditory sensory substitution device ; Research Centre in Psychology, Queen Mary, University of London, London, UK; and Sackler Centre for Consciousness Science, University of Sussex, Brighton, UK Received 5 March 2011, in revised form 31 August 2011

[4] Amedi A, Stern W, Camprodon J A, Bermpohl F, Merabet L, Rotman S, Hemond C, Meijer P, Pascual-Leone A, 2007 ``Shape conveyed by visual-to-auditory sensory substitution activates the lateral occipital complex'' Nature Neuroscience 10 687 ^ 689

[5] https://www.seeingwithsound.com/

To include : Points to expand upon :
(Related to visual processing in brain)

1. Simple , **complex cells i**nteraction, processing
2. The packaging problem - not enough neurons to have all the info of the image
3. Pyramidal model of the image vs visual processing in brain
4. Fuzzy model of human visual system - weight of different factors
5. Visual recognition - grandmother cell theory by barlow (criticism -if we lose the GM cell , then gm memory)
6. Distributed theory - criticism- how do you connect (binding problem)
7. How to use deep neural network ?
8. Visual feedback for the images - supervised vs unsupervised learning

To include in ppt :

1.
https://m.facebook.com/groups/107107546348803?view=permalink&id=575957916130428
About the NeoCortex (Jeff Hawkins 2004 book is over-detailing the subject).
This is something he left out.
While the neocortex is very generic, our behavior is not, to enforce this a lot of the input goes parthly via the Thalamus which is kind of like a genetically programmed router that makes sure that the NeoCortex gets data in such a way that:
1) we find babies cute
2) we find partners cute
3) we find fighting cute
4) we find eating fun
5) we avoid being around feces
n) …

2.
Enhance an image without loosing pixel quality : zoom more using CNN (as it has least pre processing of image and most similar to human visual system) based deep network architecture and mesh it with Photo magnifying algorithms suvh as used in Adobe Photoshop, we can get better quality pics of an image.

https://teonite.com/blog/deep-image-thanks-to-machine-learning-we-get-a-larger-image-with-a-much-better-quality-2/?utm_source=facebook&utm_medium=social&utm_campaign=blog_post_deep_image