

Деякі Аспекти Подання Візуальної Інформації Звуком для Людей з Вадами Зору

Лариса Вакал
відділ мікропроцесорної техніки
Інститут кібернетики імені В.М. Глушкова
Київ, Україна
lara.vakal@gmail.com

Євген Вакал
кафедра математичної фізики
Київський національний університет імені Тараса Шевченка
Київ, Україна
jvakal@gmail.com

Some Aspects of Visual Information Representation Via Sound for Visually Impaired People

Larysa Vakal
dept. of Microprocessor Technology
V.M. Glushkov Institute of Cybernetics
Kyiv, Ukraine
lara.vakal@gmail.com

Yevhen Vakal
dept. of Mathematical Physics
Taras Shevchenko National University of Kyiv
Kyiv, Ukraine
jvakal@gmail.com

Анотація—Розглянуто деякі аспекти сприйняття та опрацювання інформації у зоровому і слуховому сенсорних каналах людини. Проаналізовано їхнє значення для подання візуальної інформації за допомогою звуку у системах соніфікації, призначених для людей з вадами зору.

Abstract—Some aspects of information perception and procession in visual and auditory sensory channels of a person are considered. Their significance for the visual information representation via sound in sonification systems for visually impaired people is analyzed.

Ключові слова—системи соніфікації; зоровий і слуховий сенсорні канали; кодування візуальної інформації звуком

Keywords—sonification systems; visual and auditory sensory channels; visual information representation via sound

I. ВСТУП

Людам з вадами зору «бачити» дозволяють системи соніфікації, які перетворюють візуальну інформацію у звукову. Слухова система людини здатна навчитися обробляти та інтерпретувати надзвичайно складні і швидкозмінні акустичні сигнали.

У роботі проаналізовано деякі аспекти сприйняття та опрацювання інформації у зоровому і слуховому сенсорних каналах та показано їхнє значення для представлення візуальної інформації звуком у системах соніфікації, призначених для людей з вадами зору.

II. ДЕЯКІ АСПЕКТИ ПОДАННЯ ВІЗУАЛЬНИХ ДАНИХ ЗА ДОПОМОГОЮ ЗВУКУ ДЛЯ ЛЮДЕЙ З ВАДАМИ ЗОРУ

Перший аспект. Людина може опрацювати тільки обмежений обсяг інформації, що надходить до органів чуття у даний момент часу, оскільки концентрація і стійкість її уваги сильно обмежені [1]. Більшу частину інформації відфільтровує вибіркова увага. Тому кодувати звуком слід тільки критично важливу з точки зору призначення системи візуальну інформацію, щоб запобігти ризику сенсорного перевантаження слухового каналу [2, 3].

Другий аспект. При озвучуванні візуальної інформації потрібно урахувати невідповідність пропускної здатності зорового і слухового сенсорних каналів [3]. Через зір людина сприймає на два або три порядки більше інформації, ніж через слух [1]. Ця диспропорція відображається й при порівнянні числа нервових волокон: в одному зоровому нерві їх більше мільйона, у слуховому – приблизно 30000.

У зв'язку з цим рекомендується [4], по-перше, визначити які саме аспекти візуального середовища найважливіші з функціональної точки зору. Якщо, наприклад, система соніфікації призначена для розпізнавання об'єктів на зображенні, то головну увагу потрібно звернути на форму об'єктів. Якщо ж система створюється для навігації незрячих, то найважливішими аспектами візуального середовища для звукового інформування користувача є розташування і розмір перешкод на його шляху, а не їхня форма або колір. По-

друге, слід визначити як ефективно передавати цю важливу інформацію у слухову сенсорну систему. Тут потрібно враховувати основні властивості опрацювання інформації у слуховому каналі та розробляти ефективні моделі кодування візуальної інформації звуковими сигналами, наприклад, кодування напрямку перешкоди в системах навігації для незрячих за допомогою просторового звуку.

Третій аспект. Звуки, що використовуються для подання візуальної інформації, мають бути приємними для слуху людини, принаймні не дратувати її.

У даному контексті особливий інтерес становить використання музичних звуків [5]. Нейробіологи з Массачусетського технологічного інституту (США) нещодавно виявили в слуховій корі головного мозку людини популяцію нейронів, яка вибірково реагує саме на музичні звуки [6]. За словами американських учених, людина легко може відрізнити звуки музики від інших звуків навіть у тому випадку, коли вони також ритмічні.

Музичний звук – це звук тональний, який на відміну від шумів має лінійчатий дискретний частотний спектр. Така побудова музичного звуку обумовлена психофізіологічними особливостями звукового сприйняття людини, а саме: інерційністю слухового апарата та його вибірковістю в розпізнаванні звукових частот, а також здатністю пам'яті людини краще розпізнавати, засвоювати і запам'ятовувати дискретну за частотою аудіоінформацію. Вибірковість (здатність розрізнити) слухової системи людини є досить високою. У смузі частот від 0 до 16000 Гц слух людини може розрізнити до 620 градацій частот (залежно від інтенсивності звуку), при цьому приблизно 140 градацій знаходяться в діапазоні від 0 до 500 Гц.

Використання для звукового представлення візуальної інформації приємних для слуху людини музичних тонів сприятиме зручній роботі користувача з системою соніфікації упродовж тривалого часу.

Четвертий аспект. При кодуванні візуальних даних звуком доцільно використовувати переважно інтуїтивні зв'язки між візуальними характеристиками і звуками. Це значно полегшить процес навчання при роботі з системою соніфікації.

Відомо, наприклад, що для слухової системи людини інтуїтивно асоціювати вищі частоти звуку з вищим вертикальним положенням у просторі (ефект Пратта) [7]. Ця асоціація використовується, зокрема, у запропонованому в роботі [8] алгоритмі озвучування контурів об'єктів на зображенні, де звуками найбільшої частоти кодуються точки, що лежать на контурі найвище.

Проте вибір найкращої звукової аналогії для тієї або іншої візуальної характеристики ситуація часто є неоднозначним. Наприклад, виникає питання: яскравість є більш аналогічною гучності чи висоті звуку? У системі звукової візуалізації vOICE [9] яскравість пікселя відображається на гучність звуку. А в системі SIPReS [10], розробленій японськими вченими для надання незрячому користувачу наукових даних про космос, яскравість пікселя кодується частотою звуку.

Гендель у роботі [11] робить висновок, що жодна аналогія між візуальними характеристиками і звуками не є достатньою, найкраща залежить від контексту. В системах, призначених для соніфікації об'єктів на зображенні, гучніші звуки можуть відповідати крупнішим об'єктам. А в системах навігації для незрячих гучніші звуки можуть ставитися у відповідність об'єктам, які знаходяться ближче до людини (незалежно від їхніх розмірів).

Слід додати, що при озвучуванні візуальних даних використовують також характеристики звуку (тембр, часові проміжки тощо), які інтуїтивно не пов'язані з фізичними характеристиками або параметрами об'єктів, але можуть бути пов'язані з ними штучно, щоб передати більше інформації.

Отже, урахування зазначених аспектів сприятиме розробці ефективних моделей звукового подання візуальної інформації. Це допоможе людям з ослабленим зором в їхньому повсякденному житті краще сприймати візуальне середовище за допомогою слуху.

ЛІТЕРАТУРА REFERENCES

- [1] S.B. Most, B.J. Scholl, E.R. Clifford, D.J. Simons, "What you see is what you set: Sustained inattention blindness and the capture of awareness," *Psychological Review*, vol. 112, no. 1, pp. 217–242, 2005.
- [2] Á. Kristjánsson, A. Moldoveanu, Ó.I. Jóhannesson, O. Balan, S. Spagnol, V.V. Valgeirsdóttir, R. Unnthorsson, "Designing sensory-substitution devices: principles, pitfalls and potential," *Restorative Neurology and Neuroscience*, vol. 34, no. 5, pp. 769–787, 2016.
- [3] D. Osinski, M. Łukowska, D.R. Hjelm, "Wierzchon M. Colorophone 2.0: A Wearable color sonification device generating live stereo-soundscapes – design, implementation, and usability audit," *Sensors (Basel)*, vol. 21, no. 21: 7351, pp. 1–21, 2021.
- [4] M. Loomis J., R.L. Klatzky, N.A. Giudice, "Sensory substitution of vision: Importance of perceptual and cognitive processing," *Assistive Technology for Blindness and Low Vision*, Boca Raton, FL, USA: CRC Press, pp. 162–191, 2012.
- [5] L. Vakal and Y. Vakal, "Method of Encoding Contours with Musical Sound," *Applied Questions of Mathematical Modeling*, vol. 4, no. 1, pp. 50-58, August 2021.
- [6] S. Norman-Haignere, N.G. Kanwisher, J.H. McDermott, "Distinct Cortical pathways for music and speech revealed by hypothesis-free voxel decomposition," *Neuron*, vol. 88, no. 6, pp. 1281–1296, 2015.
- [7] M. Talbot, B. Cowan, "Trajectory capture in frontal plane geometry for visually impaired," *Proceedings of the 12th International Conference on Auditory Display «ICAD06»* (London, UK), pp. 16–23, 2006.
- [8] L. Vakal, Y. Vakal, "Objects' contours sonification for visually impaired people," in *Information Technologies and Computer Modelling: International Scientific and Practical Conference ITCM-2023, Precarpath. Nation. Univ., Ivano-Frankivsk*, pp. 10–11, 2023.
- [9] P.B.L. Meijer, "An Experimental System for Auditory Image Representations," *IEEE Transactions on Biomedical Engineering*, vol. 39, no. 2, pp. 112-121, February 1992.
- [10] S. Uno, Y. Suzuki, T. Watanabe, M. Matsumoto, Y. Wang, "Sound-based image and position recognition system 'SIPReS'," *Proceedings of the 24th International Conference on Auditory Display (ICAD 2018)*, Michigan Technological University, USA, pp. 67–72, 2018.
- [11] S. Handel, "Space is to time as vision is to audition: seductive but misleading," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 14, no. 2, pp. 315–317, 1988.