

УДК 004.5

Крак Ю.В., д.ф.-м.н., проф.,  
Коваль Ю.В., асистент,  
Тернов А.С., к.т.н.

### До розробки інтерактивного інтерфейсу моделювання та розпізнавання жестової інформації

Київський національний університет імені  
Тараса Шевченка, 83000, м. Київ, пр-т.  
Глушкова 4д,  
e-mail: kafedraTK@unicyb.kiev.ua

Iu. V. Krak, Prof.,  
Iu. V. Koval, Assistant,  
A. S. Ternov, PhD.

### To creation of interactive interface for sign information modeling and recognition

Taras Shevchenko National University of Kyiv,  
83000, Kyiv, Glushkova ave., 4d,

e-mail: kafedraTK@unicyb.kiev.ua

*В статті запропонована загальна архітектура системи для інтерактивно жестового інтерфейсу комунікації людина – комп'ютер. Запропоновано вимоги та принципи до функціонування такої системи.*

*Ключові слова: жестова мова, інтерфейс жестової мови.*

*A human – computer interface was discussed in the article. A common architecture of sign language human – computer interface system was discussed and proposed. The main requirements to the system, main principles of architecture construction and their structure were described. Architecture of sign language interface system divided in three levels: data adaptation level, functional modules level and design making, synchronisation and control level. Parallel independent execution of functional modules in main system and their asynchronous control were proposed. System time architecture was discussed and two different solutions proposed. Particular design of what solution to chose was postponed for practical realisation time. Size and bandwidth of human – computer interfaces also discussed. Large amount of data was pointed as crucial problem for such interface realization. Sign and alphabet data representation was discussed as algorithm selection reason.*

*Key Words: sign language, sign language interface.*

Статтю представив д.ф.-м.н., проф. Анісімов А.В.

#### 1. Вступ.

Розвиток комп'ютерів породжує розвиток їх інтерфейсів, тобто засобів комунікації. Особливе місце серед таких інтерфейсів посідають інтерфейси комунікації з людиною [1]. Традиційна клавіатура частково поступилася місцем мишкам, джойстікам, тачпадам, тачскрінам. Екран поки що залишається незамінним, хоча з'являються нові пристрої відображення інформації як то проектори. Друковані інформація та інтерфейси для її створення, незважаючи на їх важливу роль, також поступаються місцем новішим засобам. Людина має зір, слух, голос, нюх, смак, відчуття дотику, тепла, гравітації, болю. Майбутні інтерфейси

комп'ютера мають спиратися на ці можливості. Звуковий інтерфейс на сьогодні присутній, хоча може ще не зовсім досконалий. Інтерфейси для решти відчуттів або не розроблені або їх розробки поки що не вийшли за межі лабораторій. Проте певну тенденцію розвитку інтерфейсів вже можливо помітити. Якщо перший інтерфейс – клавіатура – був абсолютно надійним та правильним, то вже мишка та джойстик дозволяли введення інформації з певною точністю. При цьому слід зауважити, що вимога до зменшення точності йшла від користувачів. Спозиціонувати курсор миші з точністю до пікселя таки досить важко. В подальшому позиціонування курсору стало

взагалі непотрібно, бо змістовна частина покладалася на форму та напрямок руху. Точність руху для введення інформації від цього зменшилася. Між точністю руху для введення інформації та кількістю значень, що можуть бути введені, існує кількісна залежність – чим менша точність руху, тим менше значень можемо ввести. Якщо простір всіх можливих рухів розбити на класи еквівалентності за типом руху, то отримуємо кількість значень. Чим менша точність – тим більше елементів в класі. Це у випадку чіткої межі між формами та напрямками руху. Проте зрозуміло, що присутня певна, доволі низька, точність визначення цих форм та напрямків. Внаслідок цього виникають форми та рухи, які не можуть бути віднесені до жодного класу, а від так і задати якое значення. На додачу до цієї особливості, маємо властивість про збільшення обчислювального ресурсу, що необхідно задіяти для визначення значення, що вводиться з інтерфейсу. На підтвердження цього факту, порівняємо кількість даних, що передаються з інтерфейсу в комп'ютер для обчислення значення:

1. Традиційна клавіатура – 1 Б
2. РС-клавіатура – 1 скан код
3. мишка або тачпад – не менше 300 Б за секунду.
4. звуковий інтерфейс – не менше 3кБ за секунду
5. відеоінтерфейс – не менше 1 мБ за секунду.

Як наслідок такого зростання, зростає і обчислювальний ресурс для визначення значення, що вводиться, з інтерфейсу.

Подібне становище і для інтерфейсів, за допомогою яких комп'ютер передає інформацію людині. Текстовий термінал має найменше додаткових видатків на подання інформації. Відео термінал застосовує потужності знакогенератора. Графічне зображення у графічному терміналі генерує графічний адаптер за потужністю практично рівний головному комп'ютеру. Звуковий вивід, в той же час, вимагає потужності доволі помірної та пропорційної співвідношенню кількості графічної та звукової інформації.

## 2. Постановка задачі.

Чим більше обробляється інформації для введення або виведення даних, тим простіше виникає помилка з одного боку, і тим більше різних методів, що можуть бути застосовані до такої обробки. Так, наприклад, для розпізнавання знаків жестової мови можуть бути застосовані різні підходи і методи [2-4]. В майбутньому кількість цих методів може зрости. Кожен із запропонованих методів має свої недоліки та переваги. Тому загальний принцип роботи систем введення інформації обробкою потоку даних має змінитися від вибору алгоритму, який працює для всіх випадків до системи конкурентного розпізнавання декількома алгоритмами. Виходячи з цього формулюємо постановку задачі: розробити архітектуру системи для узгодження роботи різнорідних модулів розпізнавання рухів жестової мови.

## 3. Метод вирішення задачі.

Враховуючи складність системи та можливість її подальшого розширення

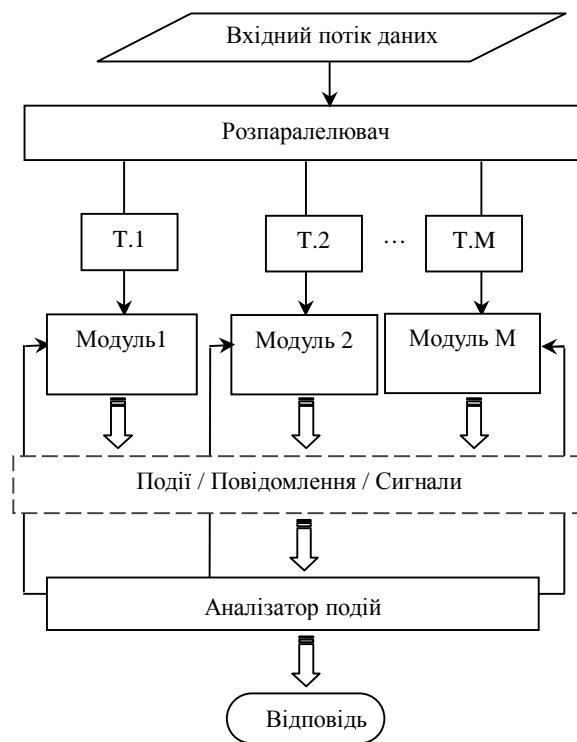


Рис. Загальна архітектура системи

функціональними модулями, загальна архітектура системи пропонується у вигляді, який подано наступною схемою (див. рис.).

Архітектурно системи розподілена на три частини:

- 1) частина системи, яка займається підготовкою даних для подальшої обробки і аналізу;
- 2) функціональні та логічні модулі;
- 3) система прийняття рішень, керування та синхронізації модулів.

Завдання, які покладені на блок «розпаралелювач», зводяться до одночасного і миттєвого (в ідеальному випадку) надання всім функціональним модулям системи даних для обробки.

Додаткові перетворювачі  $T, I, \overline{M}$  призначені для підготовки вхідних даних (їх попередня обробка) для відповідного функціонального модуля системи з метою узгодження їх з вимогами модуля, якщо такі визначено.

Функціональні модулі системи – певні незалежні обробники вхідних даних мета яких визначити інформацію представлену даними. За визначенню інформацією формуються повідомлення, які направляються в систему обробки повідомлень та прийняття рішень. В свою чергу система прийняття рішень, при необхідності, надсилає повідомлення функціональним модулям з метою узгодження та синхронізації їх роботи. Кількість модулів може варіюватись в залежності від набору алгоритмів обробки даних які є в наявності і обмежена лише технічними можливостями.

Аналізатор подій приймає рішення про розпізнавання інформаційної одиниці вхідних даних. Для звичайного інтерфейсу: клавіатура-термінал - запропонували досить природню позицію, що одиницею введеної інформації є буква. Мінімізувавши кількість вхідних або вихідних інформаційних одиниць вдалося підвищити надійність інтерфейсу для уникнення помилок технічного характеру. На противагу європейській традиції абетки існує світова традиція змістовних знаків. Її частиною є традиція ієрогліфів, хоча, безумовно, традиція знаків набагато ширша. Незважаючи на обов'язковість в сучасній

освіті вивчення традиції абетки, з раннього дитинства людям прищеплюють традицію знаків. Знак, як означення абстрактного поняття глибоко увійшов в нашу культуру і використовується на всіх рівнях, включно з юридично відповідальними місцями, як наприклад, правила дорожнього руху та техніка безпеки.

Жестова мова[5,6] на сьогодні має в собі як елементи знаків, так і елементи абетки. Виходячи з цього поєднання стає зрозумілою наявність принципово різних алгоритмів розпізнавання знаків жестової мови. На долю аналізатора подій припало узгодити різноманітну поведінку алгоритмів, на яку вже вказано. Для узгодження роботи алгоритмів має бути введена єдина система часу та подій. Завершення будь-якого алгоритму не обов'язкове, а також, алгоритм (модуль) може бути примусово зупинений аналізатором подій виходячи з рішень, прийнятих іншими алгоритмами (модулями).

Одним із способів обліку та визначення часу в системі є прийняття абсолютного часу, під яким будемо розуміти час так, як його сприймає людство в цілому: є один годинник з яким всі решта годинників синхронізуються. Синхронізація годинників стала на сьогодні домінуючим способом визначення часу в усіх комп'ютерних системах від end-user до найбільш потужних серверів. Тоді кожна подія, що відбувається в системі повинна мати мітку часу. Узгодженість дій частин системи в часі, знову ж таки, забезпечується аналізатором подій.

Іншим можливим способом обліку часу є використання часу як міри надходження даних на обробку. Тоді поняття час взагалі можна замінити порядковим номером одиниці вхідних даних. Цей підхід суттєво спрощує орієнтування в часі обробників послідовностей даних, проте має один суттєвий недолік. Вхідні данні можуть бути неоднорідними. Це пов'язано з достатнім різноманіттям пристроїв введення початкових даних як, наприклад, фотокамери, спеціалізовані камери зчитування рухів, тачпади, тачскрін, мишки, джойстики або інші пристрої, які практично

поки що не реалізовані, але які можуть і будуть реалізовані для визначення руху (наприклад 3D мишка).

Остаточне рішення, що до обліку часу буде прийнято при практичній реалізації системи і можливо перероблене в майбутніх реалізаціях.

#### 4. Висновки та подальші напрямки роботи.

В результаті проведених досліджень була розроблена та запропонована загальна

архітектура системи аналізу жестового мовлення в реальному часі та принципи взаємодії модулів. Дана система інтегрується до систем повноцінної підтримки інтерактивної взаємодії людини-носія жестової мови з комп'ютером [7,8].

Подальші дослідження спрямовані на реалізацію запропонованих функціональних модулів та модулів керування, створення загальної системи розпізнавання жестової мови.

#### Список використаних джерел

1. Летичевський О.А. В.М. Глушков і сучасна інформатика (від теорії автоматів до когнітивних архітектур) / О.А. Летичевський // Вісн. НАН України. – 2013. – № 8. – С. 21-33.
2. Vapnik V.N. Statistical Learning Theory. / V.N.Vapnik // New York: Wiley, 1998
3. Wang, X. An HOG-LBP Human Detector with Partial Occlusion Handling /X.Wang, T.X.Han // IEEE Computer Engineering. – 2009. – P. 32-39.
4. Kirichenko N.F. Analysis and synthesis of signal classification systems by means of perturbation of pseudoinverse and projection operations / N.F.Kirichenko, G.I.Kudin // Cybernetics and Systems Analysis. – 2009. – №3. – P. 47-57.
5. Stokoe W.C. Sign Language Structure: An Outline of the Visual Communication Systems of the American Deaf / W.C.Stokoe // Studies in Linguistics. – 1960. – №8. – 78 p.
6. Sing Languages. Ed. by D. Brentari. Cambridge: Cambridge University Press, 2010. – 714 p.
7. Кривонос Ю.Г. Формалізація жестоутворення за допомогою трьохмірної моделі людини для української жестової мови / Ю.Г.Кривонос, Ю.В.Крак, О.В.Бармак, Г.М.Єфімов // Проблеми програмування. – 2012, №2-3. – С. 398-405.
8. Крак Ю.В. Розробка архітектури та основних інструментів комп'ютерної анімації для побудови системи синтезу жестової мови. / Ю.В.Крак, А.С.Тернов, М.П.Лісняк // Штучний інтелект. – 2013. – №3(61). – С.147-153.

#### References

1. LETYCHEVSKYY O.A. (2013) V.M. Glushkov and modern science (from automata theory to cognitive architectures). Visn. NAS of Ukraine. 8. p. 21-33.
2. VAPNIK V.N. (1998) *Statistical Learning Theory*. New York: Wiley,
3. WANG X., HAN T.X. (2009) *An HOG-LBP Human Detector with Partial Occlusion Handling*. IEEE Computer Engineering. p.32-39.
4. KIRICHENKO N.F., KUDIN G.I. (2009) *Analysis and synthesis of signal classification systems by means of perturbation of pseudoinverse and projection operations*. Cybernetics and Systems Analysis. 3. p.47-57.
5. STOKOE W.C. (1960) *Sign Language Structure: An Outline of the Visual Communication Systems of the American Deaf*. Studies in Linguistics. 8. p.78.
6. BRENTARI D. (ed.) (2010) *Sing Languages*. Cambridge: Cambridge University Press,
7. KRYVONOS I.U.G., KRAK I.U.V., BARMAK G.M., EFIMOV A.V. (2012) *Formalizing of gestures creation using three-dimensional models for human Ukrainian Sign Language*. Problems of Programming. 2-3. p.398-405.
8. KRAK I.U.V., TERNOV A.S., LISNYAK M.P. (2013) *Development of architecture and basic tools for building computer animation synthesis system of sign language*. Artificial Intelligence. 3(61). p.147-153.

Надійшла до редколегії 29.10.14