

УДК 004.588: 376.356

Метод пошуку елементів відеозображення у тренажері для навчання мові жестів людей з вадами слуху

М. В. Давидов, Ю. В. Нікольський, О. В. Пасічник, С. В. Тиханський

*Національний університет "Львівська політехніка",
Львівська НВК "Школа-гімназія "Сихівська""*, Україна

Розглядається комплекс задач із створення спеціального тренажера для навчання мові жестів. Запропоновано метод розпізнавання обличчя та траєкторії руху долоней, в основу якого покладено сучасні технології комп'ютерного зору. Метод використовує AdaBoost, нейронні мережі та евристичний алгоритм пошуку прямокутників.

Ключові слова: *комп'ютерний тренажер, мова жестів, дистанційне навчання, ідентифікація координат, відеозображення, класифікатор, нейронна мережа.*

Рассматривается комплекс задач по созданию специального тренажера для обучения языку жестов. Предложен метод распознавания лица и траектории движения ладоней, в основу которого положены современные технологии компьютерного зрения. Метод использует AdaBoost, нейронные сети и эвристический алгоритм поиска прямоугольников.

Ключові слова: *комп'ютерний тренажер, язык жестов, дистанционное обучение, идентификация координат, видеоизображение, классификатор, нейронная сеть.*

A tool for sign language distant learning is developed. Method for face and hands tracking for sign language teach system implementation is proposed. Method takes benefits from applying AdaBoost method, neural networks and heuristic algorithm for rectangular area tracking.

Keywords: *computer simulator, sign language, distance learning, identify the coordinates of a video image, a classifier, neural network.*

1. Загальна постановка задачі та її актуальність

Проблеми, що виникають в процесі соціальної адаптації людей, які мають вади слуху, генерують низку специфічних завдань, які є унікальними в плані їх постановки та в підходах до їхнього вирішення. Основною проблемою, яка породжує багато інших, є проблема повноцінного та адекватного спілкування з такою категорією осіб. Таке спілкування вимагає володіння спеціальними засобами, основним з яких є жести мови. Один із шляхів залучення людей з вадами слуху до активної участі у повноцінному житті суспільства полягає в оперативному наданні їм актуальної інформації. Для цього люди з вадами слуху повинні приймати участь у заходах, під час яких відбувається обмін такою інформацією: дискусіях, диспутах, конференціях. Основною проблемою, яку доводиться вирішувати організаторам таких заходів, є пошук та залучення сурдоперекладачів, які забезпечать спілкування людям з вадами слуху.

Навчання мові жестів породжує ще кілька проблем. По-перше, проблему підготовки та підвищення кваліфікації викладача, який вчить мові жестів, та методики викладання. По-друге, проблема залучення тренера та наставника для навчання мові жестів такої людини, що втратила слух, або того, у кого виникла потреба в оволодінні мовою жестів для спілкування з особами, що втратили слух внаслідок хвороб або травм.

За даними Всесвітньої організації здоров'я, до 5% усіх людей страждають різними порушеннями слуху, а 0,5% страждають порушеннями слуху у тяжкій формі. В Україні чисельність такої специфічної категорії оцінюється у 0,5 млн. осіб, а, з врахуванням постійного її зростання, чисельність групи, яка потребує спеціальних засобів навчання, досягає 1,5 млн. осіб. Зрозуміло, що організація та забезпечення індивідуального навчання, консультування та методичного супроводу такої категорії осіб силами вчителів спеціалізованих шкіл є завданням, вирішення якого традиційними способами є нереальним. Проблема має загальнодержавний, соціально-орієнтований характер, адже в такому становищі може виявитися значна кількість людей. На часі розроблення сучасних комп'ютеризованих засобів, які допоможуть навчати, коректувати процес навчання під індивідуальні особливості учня, дозволяти самостійно вдосконалювати мовні навички.

Для підвищення ефективності навчання мові жестів пропонується залучити сучасні інформаційні технології, а саме, створити спеціальний програмно-комп'ютерний тренажер, який допоможе масовому навчанню мові жестів, викладачів мови жестів, а також створити спеціальні засоби для самостійного вивчення та покращення володіння українською мовою жестів широкою категорією людей, які мають проблеми із слухом. Розроблені авторами програмні засоби [1-4] демонструвалися на міжнародній виставці ЦеБІТ у 2006, 2007 та 2008 рр. (м. Ганновер, Німеччина) у рамках стенду Міністерства освіти та науки України та отримали схвальні відгуки учасників. Ці засоби заклали основу для вирішення проблеми навчання дітей з вадами слуху у неспеціалізованих загальноосвітніх навчальних закладах.

2. Сучасний стан проблеми

Зважаючи на актуальність поставленої проблеми, у світі ведуться активні дослідження із впровадження сучасних інформаційних технологій у процес навчання мові жестів. На світовому ринку існують комерційні розроблення [5-7], що дозволяють вирішувати певні завдання навчання мові жестів, проте комплексного засобу, який покриватиме всі інформаційні потреби людей з пониженим слухом з навчання мові жестів, немає. В Україні відсутні аналоги пропонованої системи, що базуються на українській жестовій мові. Використання інформаційних технологій у сучасному процесі навчання мові жестів у нашій країні здебільшого обмежується використанням відеоматеріалів. В Україні навчання мові жестів на сьогодні має свої особливості. Вивчення жестової мови дітьми відбувається у спеціалізованих школах для дітей з вадами слуху під керівництвом вчителів. При цьому використовують різні технічні засоби у вигляді спеціальних відеопідручників. Люди, які втратили слух навчаються під опікою Українського товариства глухих, Об'єднання нечуючих педагогів, добротинних організацій, спеціалізованих курсів та шкіл, де використовують традиційні некомп'ютеризовані технології. Підготовка сурдоперекладачів здійснюється у вищих навчальних закладах, зокрема, Інституті корекційної педагогіки та психології НПУ імені М.П.Драгоманова, Відкритому міжнародному університеті розвитку людини „Україна” тощо, з використанням традиційних навчальних технологій.

У масштабах країни навчання традиційно відбувається з участю викладача та тренера. У процесі навчання використовують відео та графічні матеріали, але не застосовують спеціалізовані програмно-комп'ютерні системи підтримки процесу навчання, контролю за засвоєнням, коректування цього процесу з метою виправлення помилок тощо. Розроблення проекту тренажера з україномовним наповненням не має аналогів у світі.

Ще одним суттєвим недоліком застосовуваної нині системи навчання є те, що відсутня система дистанційного навчання. Таке навчання можна реалізувати в реальному часі із використанням найновіших методичних матеріалів, які можна, в разі реалізації проекту, використовувати та поширювати в режимі дистанційної освіти. Дистанційна складова проекту дозволить також підвищувати кваліфікацію вчителів з мови жестів та фахових сурдоперекладачів, залучати сурдоперекладачів до участі в певних заходах у режимі телеконференцій.

Розроблюваний комплекс засобів для навчання мові жестів має на меті усунення недоліків сучасного підходу до цієї підготовки на основі використання переваг сучасних інформаційних технологій.

Розроблена комплексна система навчання мові жестів є платформою, на якій планується створити цілий комплекс спеціальних засобів додаванням до нього нових програмних інструментів [1-2]. При цьому розроблена на сьогоднішній день система вже містить необхідний мінімум таких засобів, які забезпечують навчання людей мові жестів, і може бути розширеною у подальшому додатковими засобами. Програмний продукт дозволяє вирішувати такі завдання навчання жестовій мові: пошук жестів за словником жестової мови, навчання та самоконтроль за словником, дистанційна робота із вчителем у режимі відеоконференції.

3. Постановка задачі

Люди з пониженим слухом спілкуються спеціальною жестовою мовою, складовими якої є набори жестів. Спілкування з особами, які користуються жестовою мовою, вимагає залучення сурдоперекладачів. Таких фахівців є надзвичайно мало, і тому спілкування двох осіб, одна з яких володіє жестовою мовою, а друга – ні, є суспільною проблемою, яка обмежує можливості людей з пониженим слухом. Ця проблема постає у процесі дискусій, передаванні великого обсягу інформації під час конференцій, виступів з промовами, інших громадських заходів, навчання у спеціалізованій школі або в разі вивчення жестової мови людьми, які втратили слух.

Усе перераховане спричиняє необхідність створення спеціальних засобів автоматизації обміну інформацією з людьми з послабленим слухом. Сучасні дослідження спрямовані на створення комп'ютерно-програмних систем, у яких використовується для спілкування українська мова та врахована специфіка мови жестів з україномовною семантикою. Завданням досліджень є створення системи спілкування із вказаними групами людей шляхом використання сучасних інформаційних технологій. Один зі шляхів вирішення проблеми спілкування полягає у створенні систем підтримки засобів комунікації з людьми з пониженим слухом. Рішенням цього завдання є програмно-апаратний комплекс, що містить у собі засоби для вирішення проблем навчання

сурдоперекладу, у який мають входити такі компоненти: словник мови жестів, тренажер мови жестів, засоби для дистанційного навчання мови жестів, засоби аналізу системи візуальних сигналів, перетворення таких сигналів у текстові повідомлення на екрані комп'ютера. Важливою вимогою до елементів цього комплексу є можливість їхнього використання у реальному часі. Нині у різних країнах світу виконуються окремі дослідження, які зорієнтовані на вирішення окремих завдань, однак їх постановки та рішення як цілісного комплексу немає. В Україні відсутні аналоги пропонованим дослідженням, а розроблення системи із врахуванням специфіки україномовного контенту пропонується вперше.

Розроблення засобів дистанційного навчання жестовій мові вимагає створення засобів автоматизованої перевірки правильності показу жестів учнем. Для розв'язання цієї задачі необхідно в першу чергу ідентифікувати координати обличчя та долоней на відеозображенні. Подальшими кроками будуть ідентифікація форми губ для визначення артикуляції та ідентифікація форми та траєкторії долоней для визначення показаного жесту.

Задачею статті є пошук обличчя та долоней на відеозображенні.

4. Метод пошуку обличчя та долоней користувача на відеозображенні

Пошук пошуку обличчя та долоней користувача виконується на кадрах відео, яке надходить з веб-камери. Веб-камера фронтально знімає користувача так, що обличчя і долоні одночасно знаходяться в полі зору камери.

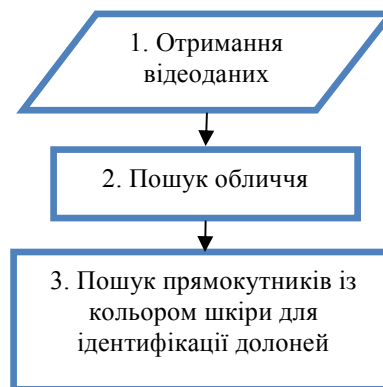


Рис.1. Схема процесу визначення положення обличчя та положення долоней на відеозображенні.

Ідентифікація обличчя та долоней виконується за такою схемою (рис. 1):

- 1) Отримання відеоданих;
- 2) Пошук обличчя;
- 3) Пошук прямокутників із кольором шкіри;
- 4) Визначення траєкторії руху долоней.

Отримання відеоданих виконується з частотою кадрів 10 кадрів/с та роздільною здатністю 320x240 пікселів.

Для пошуку обличчя використано метод характеристик Хаара. Характеристики Хаара обчислюють як різницю між сумами інтенсивності пікселів в межах двох прямокутників, що знаходяться всередині зображення, яке

класифікується. Існують варіанти використання характеристик Хаара для кількості прямокутників більшої від двох. У такому випадку для кожного прямокутника обчислюється його інтегральна характеристика, і на основі цих характеристик будується класифікатор.

Найпоширенішим є використання інтегральних прямокутників, сторони яких або паралельні сторонам зображення, або знаходяться під кутом 45^0 до них. Хоча для точнішого класифікатора логічно використати довільні прямокутники, проте, це не є раціональним з точки зору швидкодії.

Однією з переваг характеристики Хаара є можливість швидкого обчислення необхідних параметрів. При потребі обчислити суму яскравостей пікселів використано попередньо обчислені інтегральні таблиці, в які записано суму значень яскравості пікселів, що знаходяться зліва і вище даного пікселя. У такий спосіб зменшується обчислювальна складність алгоритму, а класифікатор, побудований за таким принципом може бути використаний у системах реального часу.

Для побудови ефективного класифікатора використано трірівневу систему класифікаторів. Класифікатор першого рівня (слабкий класифікатор) є деревом рішень, кожен елемент якого складається з опису характеристик Хаара і значення точності, на яку може відхилитися зображення, яке класифікується, як еталонне. Дуже поширений варіант, коли таке дерево складається лише з одного елемента – кореня. На виході класифікатор повертає значення 1 або -1 залежно від того, яку ділянку зображення він класифікував: схожу на еталонне чи ні.

Класифікатор другого рівня (сильний класифікатор) використовує кілька слабких класифікаторів для покращення якості класифікації. Сильним класифікатором другого рівня взято класифікатор AdaBoost. При побудові класифікатора другого рівня на основі AdaBoost будується нелінійний сильний класифікатор $H_M(x)$ як лінійна комбінація M слабких класифікаторів

першого рівня у вигляді $H_M(x) = \left(\sum_{m=1}^M \alpha_m h_m(x) \right) / \sum_{m=1}^M \alpha_m$, де x – зображення, що

класифікується, $h_m(x) \in \{-1; +1\}$ – M слабких класифікаторів, $\alpha_m \geq 0$ – вагові коефіцієнти, $\sum_{m=1}^M \alpha_m$ – нормуючий множник. Використано дискретні слабкі класифікатори $h_m(x)$, які приймають значення $\{-1; +1\}$.

Процедура навчання сильного класифікатора AdaBoost полягає у знаходженні найкращих слабких класифікаторів $h_m(x)$ та найкращої комбінації вагових коефіцієнтів α_m .

Для навчання задається набір прикладів $\{(x_1, y_1), \dots, (x_N, y_N)\}$, де x_i – вектор атрибутів, $y_i \in \{-1; +1\}$ – еталонне значення для класифікації, M – необхідна кількість слабких класифікаторів. Вагові коефіцієнти $[w_1, \dots, w_N]$ позначають складність класифікації відповідного прикладу. Що приклад складніший для розпізнавання, то більше уваги приділяється йому на наступній

ітерації. Перед початком роботи алгоритму всім прикладам надано однаковий коефіцієнт складності.

Класифікація виконується за M ітерацій. Кожна ітерація складається з таких дій.

1. Знаходження оптимального слабкого класифікатора $h_m(x)$ з врахуванням вже побудованого сильного класифікатора $H_M(x)$.
2. Обчислення коефіцієнту α_m так, щоб мінімізувати помилку класифікації.
3. Обчислення вагових коефіцієнтів $[w_1, w_2, \dots, w_N]$ так, щоб приклади, які дають більшу помилку класифікації, мали більший ваговий коефіцієнт.

В результаті навчання отримаємо набір слабких класифікаторів $h_m(x)$ та набір відповідних значень коефіцієнтів α_m .

Класифікатор третього рівня (каскад сильних класифікаторів) об'єднує кілька класифікаторів другого рівня логічною функцією "І". Якщо принаймні один з класифікаторів не класифікує зображення як еталонне, то зображення у подальшому не перевіряється. На практиці спочатку перевіряються класифікатори, які можуть бути обчислені з найменшими витратами часу. Каскад класифікаторів дозволяє поєднати швидкість класифікації і точність класифікації. При цьому більшість зображень буде відкинута вже на перших класифікаторах. Для класифікації обличчя за даним алгоритмом використана програмна бібліотека OpenCV.

Для пошуку прямокутників із кольором шкіри кожному пікселю зображення ставиться у відповідність число, яке визначає подібність околу пікселя до шкіри. Для визначення належності околу шкірі використано нейромережвий класифікатор. Нейронна мережа має 15 входів, один призований шар та один вихід. Для формування входу нейронної мережі розглядається п'ять пікселів, які створюють хрестоподібну область (рис. 2). Характеристиками кожного пікселя є компоненти кольору у форматі RGB.

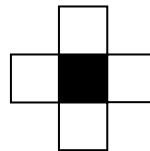


Рис.2. Хрестоподібний окіл для вибору характеристик зображення. Чорним кольором позначено піксел зображення, окіл якого розглядається.

Значення виходу нейронної мережі, яке приймає значення більше від 0.5, вважається сигналом про належність певної області шкірі людини. Відповідно, значення, яке менше або рівне 0.5, вважається сигналом того, що область не є шкірою.

Використання нейронної мережі дозволяє використати певні дані про текстуру області та покращує виділення долоней користувача порівняно із методом моделювання шкіри лише за кольором одного пікселя [8].

Коли кожному пікселю зображення поставлена у відповідність числова характеристика подібності до шкіри, використовується евристичний алгоритм пошуку прямокутників. Параметрами алгоритму є мінімальні ширина та висота прямокутника L_{\min} та мінімальне середнє значення ступеня подібності до шкіри пікселів всередині прямокутника P_{\min} .

Алгоритм знаходження прямокутників, які можуть містити долоні користувача.

1. Створити список прямокутників *Rectangles*, який початково містить один прямокутник – знайдене обличчя людини.
2. Обрати прямокутник розміру $L_{\min} \times L_{\min}$, який не перетинається із жодним з прямокутників *Rectangles* та має найбільше середнє значення ступеня подібності до шкіри пікселів всередині нього. Якщо це середнє значення менше за P_{\min} , то закінчити пошук.
3. Розглянути варіанти збільшення прямокутника вгору, вниз, вліво, вправо. Обрати напрямок збільшення, який в результаті дає найбільше середнє значення ступеня подібності до шкіри пікселів, що попадуть до прямокутника.
4. Якщо жоден з напрямків розширення не задовольняє умові на мінімальне значення P_{\min} , додати прямокутник до списку *Rectangles* та перейти до кроку 2. Інакше розширити прямокутник у вибраній напрямком та перейти до кроку 3.

Параметри P_{\min} та L_{\min} обирають на етапі налаштування системи та залежать від віддалі до доповідача та параметрів експозиції.

Вважається, що два найбільші прямокутники, крім обличчя є долонями користувача, які використовують для подальшого опрацювання.

Подальше оброблення зон із долонями полягає у порівнянні форми долоні з еталонними та визначенні правильності показу жеста [9].

5. Результати тестування

Результати тестування системи на зображеннях, отриманих в складних умовах освітлення показали працездатність запропонованого підходу та невелику кількість фальшивих спрацювань алгоритму. На рис. 3 наведені результати опрацювання зображення. За допомогою підбору параметрів алгоритму забезпечено правильне виділення усіх положень долоні в кадрі. У цьому випадку в середньому один раз на 3-4 с виникали хибні спрацювання алгоритму, які усунуто на етапі порівняння з еталоном. Хибні спрацювання були спричинені подібністю кольору светра та кольору шкіри в наведеному прикладі. Час опрацювання одного кадру розміру 320x240 складає 120 мс на комп'ютері Pentium 2 GHz, що є прийнятним для опрацювання в реальному часі.

6. Висновки

В процесі розробки тренажерного комплексу було створено програмний засіб, який дозволяє здійснити автоматизацію процесу навчання мові жестів за допомогою сучасних інформаційних технологій. Засіб вирішує важливу соціальну проблему дистанційного навчання жестовій мові. Невибагливість до апаратного забезпечення дозволяє використовувати програмний засіб на всіх сучасних комп'ютерах. Для самоконтролю достатньо простої веб-камери.

Розроблені засоби попереднього пошуку положення долоней користувача значно пришвидшує процес ідентифікації за методом еталону, що дозволяє в реальному часі ідентифікувати жести доповідача.

В подальшому планується додати повноцінне розпізнавання жестової мови із розпізнаванням міміки, артикуляції та жестикулювання з метою створення тренажера жестової мови. Також ведуться роботи із розробки повноцінної системи трансляції жестової мови у текст.

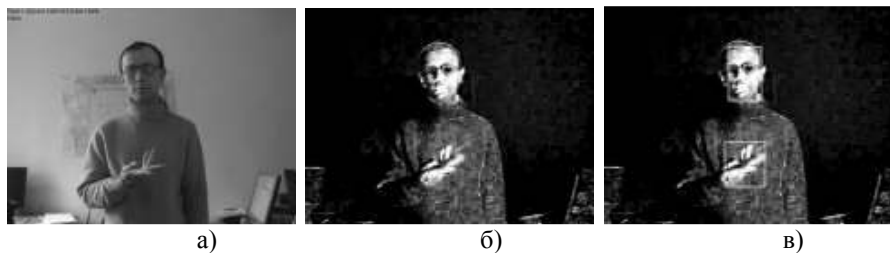


Рис. 3. Етапи опрацювання зображення: (а) – початкове зображення; (б) – результат застосування нейронної мережі для визначення ділянок кольору шкіри; (в) – результат пошуку прямокутників

ЛІТЕРАТУРА

1. Пасічник В. В. Математичне моделювання та програмна реалізація елементів тренажеру для навчання жестовій мові людей, що втратили слух / В. В. Пасічник, Ю. В. Нікольський, М. В. Давидов // Сборник трудов седьмой международной конференции «Интеллектуальный анализ информации ИАИ-2007»: 15-18 мая 2007 г. – К., 2007. – С. 56-66.
2. Пасічник В. В. Програмний тренажер для навчання мові жестів / В. В. Пасічник, Ю. В. Нікольський, М. В. Давидов // Спеціалізований тематичний додаток до загальногалузевого науково-виробничого журналу „Зв’язок”. Праці міжнародної наукової конференції „Розвиток інформаційно-комунікаційних технологій та розбудова інформаційного суспільства в Україні”: м.Ганновер, Німеччина, СеВІТ-2007, 15-21 бер. 2007 р. – К., 2007. – С. 98-106.
3. Давидов М. В. Методи та засоби опрацювання зображень реального часу для ідентифікації елементів жестової мови / М. В. Давидов, Ю. В. Нікольський // Штучний інтелект, науково-технічний журнал, державний університет інформатики і штучного інтелекту. – Донецьк, 2008. – №1. – С. 131-138.

4. Давидов М. В. Аналіз методів розпізнавання у моделях жестової мови / М. В. Давидов, О. В. Пасічник, Ю. В. Нікольський // Східно-Європейський журнал передових технологій. – Харків, 2008. – №4/2 (34). – С. 57-61.
5. Zahedi M. Appearance-Based Recognition of Words in American Sign Language / M. Zahedi, D. Keysers, H. Ney // Pattern Recognition and Image Analysis. – МАІК, 2005. – P. 511-519.
6. Zahedi M. Combination of Tangent Distance and an Image Distortion Model for Appearance-Based Sign Language Recognition / M. Zahedi, D. Keysers, T. Deselaers, H. Ney // Pattern Recognition. – Springer, Berlin, 2005. – P. 401-408.
7. Dreuw P. Speech Recognition Techniques for a Sign Language Recognition System / P. Dreuw, D. Rybach, T. Deselaers, M. Zahedi, H. Ney // Interspeech. – August, 2007. – Antwerp, Belgium, 2007. – P. 2513-2516.
8. Raja Y. Tracking and Segmenting People in Varying Lighting Conditions Using Colour / Y. Raja, S. J. McKenna, S. Gong // Proceedings of the 3rd. International Conference on Face & Gesture Recognition. – 1998. – p. 228 – 233.
9. Давидов М. В. Вибір ефективного методу опрацювання зображень на основі еталону для ідентифікації елементів жестової мови / М. В. Давидов, О. В. Пасічник, Ю. В. Нікольський // Автоматизированные системы управления и приборы автоматики, Всеукраинский межведомственный научно-технический сборник. —2008. — №139. — С. 59—68.

Надійшла у першій редакції 27.07.2009, в останній - 14.10.2009.

© Давидов М. В., Нікольський Ю. В., Пасічник О. В., Тиханский С. В., 2009