

Modern technológiákon alapuló otthoni felügyelő rendszer

Kalmár György¹, Balázs Péter²

¹ Szegedi Tudományegyetem, Képfeldolgozás és Számítógépes Grafika Tanszék
kalmargy@inf.u-szeged.hu

² Szegedi Tudományegyetem, Képfeldolgozás és Számítógépes Grafika Tanszék
pbalazs@inf.u-szeged.hu

Absztrakt. A betörések, rablások megelőzése vagy megakadályozása a hétköznapi ember számára csak riasztó berendezés használatával lehetséges, amely viszont drága és beszereléséhez szakemberre van szükség. Kutatásunk célja egy olyan alacsony költségű, mindenki számára elérhető, könnyen beüzemeltethető és használható, mobil alapú védelmi berendezés megalkotása, amely képes megtanulni, majd később felismerni a védelmezett terület tulajdonosának arcát, felhasználva a ma már olcsón beszerezhető új, illetve használaton kívül esett régi okostelefonokat. A kifejlesztett felügyelő rendszer jelenleg egy arcot képes megtanulni, majd azt a későbbi megjelenés esetén felismerni. Ismeretlen arc esetén MMS vagy e-mail üzenetet küld a felhasználójának az arcról készült képpel. A szakirodalomban már jól kiforrott módszereket megfelelően egy működő rendszerre integráló Android alapú alkalmazás lehetővé teszi arcról nagy felbontású képek nagyon gyors kimentését, ezzel teret adva arcfelismerő algoritmusok használatának. A cikkben a kezdeti elért eredményeinket ismertetjük.

I. BEVEZETÉS

Napjaink egyik kiemelkedő társadalmi problémája a betörések, lopások, rablások megakadályozása. Habár széles körben elérhetők otthonilag használható riasztóberendezések, ezek beszerelése sokszor túl költséges egy család számára. Továbbá egy rövid kódon múlik biztonságuk, amelyet akkor is igényelnek hangos riasztás közepette, ha a valódi tulajdonos érkezik otthonába. Megállapítható tehát, hogy ezek a rendszerek sokszor alacsony tudásúak és kényelmetlenek. Ezzel szemben egy okostelefon, amely ma már mindenki számára a hétköznapiak része, a kommunikáció és multimédia egész tárházát nyújtja egy tenyérnyi kis eszközben, magas processzor teljesítménnyel, sok memóriával, kamerával, mikrofonnal, internet hozzáféréssel, telekommunikációs funkciókkal.

Kutatásunk célja, egy olyan mindenki számára elérhető, olcsó, intelligens otthoni felügyelő rendszer kivitelezhetőségének a vizsgálata és megalkotása, amely képes megtanulni a rendszer használójának arcát, majd későbbi ismételt megjelenés esetén felismerni azt, ismeretlen személy esetén pedig riasztást küldeni MMS vagy e-mail üzenet formájában, amely tartalmaz egy képet az azonosítatlan arcról. Ezen rendszer elsősorban az elavulttá vált, sérült okostelefonok segítségével rendkívül kis költségigénnyel járulna hozzá a hétköznapi emberek biztonságához.

Ma már az okos eszközök nagyon nagy hányadán Android alapú operációs rendszer fut. Ez lehetővé teszi, hogy az egyes eszközök különbségeit a felsőbb rétegekben elrejtjük. Tehát egy elkészült, védelmi célokat biztosító alkalmazás az összes

Android alapú eszközön futni képes. Ezen megfontolásból a megvalósításához mi is Android alapú eszközt választottunk és az elkészült alkalmazást is Java nyelven, Androidos környezetben fejlesztettük ki.

II. A MEGVALÓSÍTOTT RENDSZERRŐL ÁLTALÁNOSÁGBAN

A jelen megvalósításban tárgyalt rendszer csak egy arc megtanulására képes. Ezen korlátozás kihatással van az alkalmazott módszerre. A mesterséges intelligencia megalkotásánál figyelembe kell venni, hogy a döntéshozás bináris jellegű, tehát megtanult arc vagy nem az szerepel egy képen.

Az arcfelismerés statikus képek esetén bár jól, de nem tökéletesen megoldott probléma. Esetünkben a felismerés nehézségéhez hozzájárul az is, hogy nem élhetünk azzal a feltételezéssel, hogy a felvétel az arcról szemből készül. Ezért már egy személy felismerésénél is adódnak nehézségek. Elsősorban az intraperszonális variancia, vagyis hogy egy személy saját magához képest mennyire változatos, jelenti a fő nehézséget. Akadályozó tényező továbbá, hogy esetünkben a felismerendő személyek mozognak, így csak homályos képeket készíthetünk, ha figyelembe vesszük, hogy az okostelefonok kamerája nem a legkifinomultabb optikai rendszer és nem is a gyors képkészítésre optimalizált.

A képfeldolgozási algoritmusokat az alkalmazás az OpenCV4Android nevű csomagon keresztül éri el, amely az OpenCV [1] Androidos környezetre előkészített változata. Az eljárások nagy része C++-ból fordított natív kód formájában kerül futtatásra a Java-s környezetből egy JNI (Java Native Interface) csatoláson keresztül, ezzel lényegesen felgyorsítva a végrehajtást.

A rendszerrel szemben elvárt funkciók:

- legyen képes egy olcsó okostelefonon is futni (minimális elvárás, hogy az rendelkezzen egy hátoldali kamerával, egy mikrofonnal, és képes legyen Android 4.0-s vagy újabb verziót futtatni)
- legyen képes egy arc megtanulására, ezt egy egy napos, felügyelt tanítási fázisban kell, hogy megtegye
- a tanítási fázis után képes legyen egy megjelenő arcról eldönteni, hogy az a megtanult arc (személy) vagy sem
- mindezen felismerő funkciókat a lehető leghatékonyabban és legmegbízhatóbban végezze

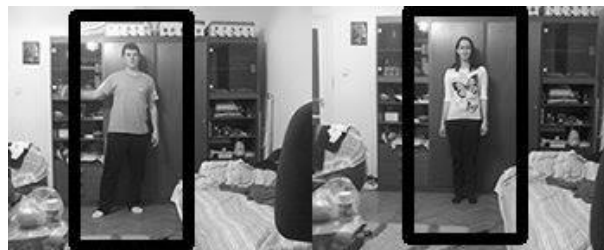
- minimalizálja a fals pozitív döntések számát, hiszen ez azt jelentené, hogy ismeretlen arcot ismertnek feltételezne
- ha az arc felismerése fizikai korlátok miatt nem is hajtható végre, akkor is próbáljon meg döntést hozni az illető kilétéről, ezt jelen megvalósításban egy hangalapú jelszóellenőrző algoritmus végzi el

A rendszer beüzemelésének kezdetekor az adott okostelefonra fel kell telepíteni az alkalmazást. Elindítva a programot, az egyből a tanulási fázisba lép. Ezen fázis során tanulja meg a később felismerni kívánt arcot, illetve a hangalapú jelszó azonosításhoz a referencia jelszót. A tanulás végeztével a rendszer egyből riasztó, felügyelő üzembe lép. Ez az üzemmód az alkalmazás leállításáig tart.

III. A RENDSZER ALAPKÖVEI

Az emberi alak megbízható és gyors felismerése az egyik, és talán a legfontosabb alappillére a felügyelő rendszernek. Ennek köszönhető, hogy a képen megjelenő személy detektálásra kerül, így a rendszer értesül arról, ha a megfigyelt térségben valaki tartózkodik. Jelenleg erre a célra, az OpenCV által is támogatott HOG leíróval működő módszert használjuk [2,3]. Ennek előnye, hogy nagyon megbízható, a fals pozitív értesítések száma minimális. Ugyanakkor igen műveletigényes, gyengébb processzoron és nagyobb felbontású képen futtatva több másodpercig is eltart az elemzés. Ennek kiküszöbölésére hoztuk létre a később tárgyalásra kerülő saját kamera kezelő osztályt, amely segítségével már a hardvertől érkező képek felbontását megfelelően alacsonyra tudjuk állítani ahhoz, hogy a HOG algoritmus szinte valós időben működhessen. Az emberi alak detektálásának működése során készült képek az 1. ábrán láthatóak.

Az emberi alak megtalálásán túl a rendszer másik elengedhetetlen eleme az arcok megbízható és gyors detektálása, tekintve, hogy a végső cél ezek megtanulása és felismerése. Még az emberi alak többnyire egy általános formának mondható, addig az arc nagyon változatos kinézetű is lehet. Az OpenCV az eddig publikált leggyorsabb és legrobosztusabb arcfelismerő eljárást támogatja. Ennek lényege, hogy úgynevezett Haar jellemzőket nyernek ki a képekből, majd ezekre építve gyenge osztályozókat hoznak létre, amelyeket egymás után fűznek (Haar cascade classifier) [4]. Az OpenCV ezen eljárásra támaszkodva, előre tanított osztályozókat biztosít, amelyek nem csak arc, de sok más objektum keresésére is alkalmasak [5]. Ennek köszönhetően választottuk az arcfelismerés módszerül azt, hogy az arcról ún. mikrojellemzőket vonunk ki és ezeket hasonlítjuk össze a később megjelenő arc jellemzőivel. A mikrojellemzők konkrétan a száj, az orr, a bal szem, a jobb szem, és a szempár, pontosabban az ezekről készült képrészletek. Az arc és arcrészek detektálására példákat a 2. és 3. ábrán láthatunk.



1. ábra: Emberi alak felismerés



2. ábra: Arc detektálása



3. ábra: Arcrészek detektálása

IV. KIEGÉSZÍTŐ FUNKCIÓK

A felügyelő rendszer alapvető funkcióinak tárgyalása után a következőkben bemutatjuk a nem szükségszerű, de nagyon hasznos kisegítő eljárások működését. Ezek közé tartozik a pehelysúlyú mozgásérzékelő, a hangalapú jelszófeldolgozó, a dinamikus arckövető algoritmus és a saját kamera kezelő osztály.

Habár az emberi alak keresés nagyon hatékony a kis méretű képen, és így a feldolgozási sebességet is magasan lehet tartani, mégis indokolatlan egy ilyen bonyolultságú algoritmust futtatni minden képre. Ha ugyanis nincs mozgás a képen, akkor biztosak lehetünk benne, hogy ember nem sétált be a megfigyelt területre. Ezt kihasználva egy mozgásérzékelő algoritmus előszűrője lehet az emberi alak keresésnek, miszerint azt csak akkor érdemes futtatni, ha történt a képen jelentős változás. Nyilvánvaló, hogy ez az algoritmus sokkal egyszerűbb, kisebb processzor és memóriaigénnyel. A szakirodalom *background subtraction*, azaz háttér kivonás néven hivatkozik az olyan algoritmusokra, amelyek célja, hogy egy képen elkülönítsék a háttérhez és az objektumhoz (előtérhez) tartozó pixeleket. A legkiemelkedőbb módszerek ezen a téren valószínűségi háttér modelleket alakítanak ki. Az általunk használt, OpenCV-ben megvalósított módszer lényege, hogy minden pixel intenzitását egy tanító képhalmaz segítségével egy normális eloszlással adja meg. Vagyis minden pixelre meghatároz egy várható értéket és szórást, és az eloszlást normálisnak feltételezi (változó paraméterű Gauss görbékkel ír le minden pixelt). Innen ered a neve, Gaussian Mixture Model, vagyis Gauss görbék keverékéből álló modell [6]. Az OpenCV-ben szereplő verzió nagyon gyors és műveletigénye is alacsony ([7]). Működése közben készült

képeket a 4. ábrán figyelhetünk meg.

Az eddig tárgyalt funkciók célja, hogy lehetővé tegyék a felügyelő rendszer számára, hogy érzékelje az emberi jelenlétet és megpróbálhassa a személy arcát felismerni. Számos okból előfordulhat azonban, hogy az arc detektálása fizikai okokból eleve lehetetlen (pl.: rossz fényviszonyok, eltakart vagy a kamerának háttal forduló arc, gyors mozgás, amely miatt lehetetlen éles képet készíteni). Ilyen esetekben is valahogy meg kellene arról bizonyosodni, hogy a detektált emberi alak ismert archoz tartozik-e.

A probléma megoldásához a hang alapú interakciót választottuk. Az azonosítás egy előre megtanult jelszó ismételt kimondásával valósul meg. A tanulási fázis elején a felhasználó egy tetszőleges jelszót mondhat be a rendszerbe, amit az felvesz és megjegyez. Ha a későbbi éles működés során az arc felismerése sikertelen lenne valamilyen okból kifolyólag, akkor a rendszer egy hangjelzés segítségével „megkéri” a felügyelt területen tartózkodó személyt, hogy ismétlje el a megtanult jelszót. Ezt felvételezi, majd kinyeri a felvételi idő beszéd részét. A referencia és az aktuálisan meghatározott beszéd régiók korrelációja alapján döntést hoz, hogy ugyanaz a szó hangzott-e el.

Az eszköz által felvett audio jel nagyon zajos lehet. Erre láthatunk példát az 5. ábrán. A jelen egyenirányítást majd szűrést hajtunk végre a következő szűrő segítségével (csúszóablakos, súlyozott, átlagoló szűrő):

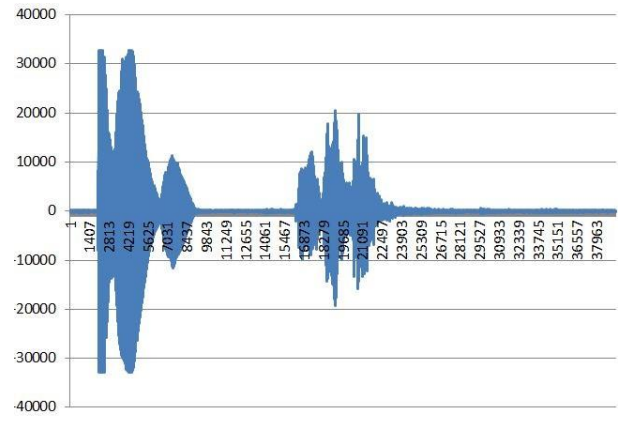
$$f(x) = \frac{1}{2n+1} \sum_{i=-n}^n (n - |i| + 1) * y(x+i), x = n, \dots, L-n$$

ahol $f(x)$ a szűrt jel az x pozícióban, $n=N/2$, ahol N a csúszóablak szélessége. L az eredeti jel hossza, $y(x)$ pedig az értéke az x pozícióban.

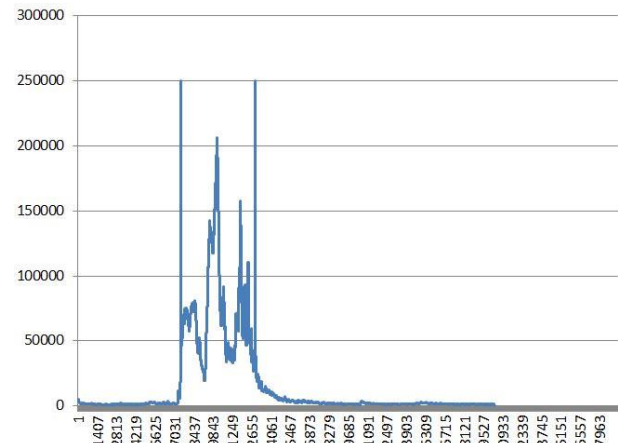
Az 5. ábrán látható jelalak elején megfigyelhető kiemelkedések a rendszer által kiadott hang miatt jelennek meg, ami mindig ugyanolyan hosszú, tehát eldobható ez a rész. Az 5. ábra jelének szűrt, kezdeti részt nem tartalmazó alakja a 6. ábrán figyelhető meg, amelyen a két kiemelkedő csúcs az algoritmus által meghatározott beszéd régió kezdő és végpontját jelöli. A beszéd régió jelen megvalósításban a leghosszabb, nem zajszinten lévő régió.



4. ábra: Mozgásérzékelés



5. ábra: Nyers hangfelvétel hullámformája



6. ábra: Szűrt jel, bejelölt beszéd régióval

Ha már rendelkezésre állnak a beszéd régiók, akkor ezeket kinyerve a referencia és az aktuálisan felvett hangból, kezdődhet az összehasonlítás. Ennek során az összehasonlítandó beszédrezt ugyanolyan hosszúra alakítjuk, mint a referencia jelszó hossza (nyújtás, majd interpoláció). Ez azért fontos, mert ugyanazt a szót többféle ritmusban is ki lehet mondani. Ha már egyforma hosszúak, normalizált keresztkorrelációval kiszámítjuk az egyezést a két hangalapú jelszó közt. Ha egy előre megadott küszöböt meghalad a hasonlóság mértéke, akkor a bediktált jelszót elfogadjuk, egyébként pedig elvetjük.

Már az előzőekben említés esett arról, hogy fontos és elengedhetetlen az arcról nagyon gyorsan, a lehető legnagyobb felbontással képeket készíteni. Ehhez szükségünk van egy olyan eljárásra, amely biztosítja, hogy ne a teljes képet kelljen átkutatni arc után. Ehhez két feltevéssel élhetünk:

1. Emberi alak fellelése után elegendő csak az emberi alak helyének felső részén keresni arc után egy adott méretű régióban.

2. Ha már találtunk arcot, valószínű, hogy a következő képen is ehhez a helyhez nagyon közel lesz az arc.

Ezeket kihasználva egy olyan módszert fejlesztettünk ki, amely nyomon tudja követni az arc helyzetét. Az eljárás egy növekvő méretű keresési ablakot használ. Az ablak mérete azzal arányosan nő, hogy hány képkocka óta nem volt arc az

adott ablakban. Ha ismét talál egy arcot, akkor az ablak középpontját áthelyezi az arc középpontjába és innen folytatja újra a növekedést, ha ismét eltűnne az arc. A 7. ábrán egy olyan sorozatot láthatunk, amelynek első képén az ablak az emberi alak detektálás után állítódott be. Majd a további képeken a mozgás hatására változó helyzetű és méretű ablak figyelhető meg.

Az alkalmazás szempontjából kritikussá vált, hogy elérhetőek legyenek a következő, kamerát érintő funkciók:

1. Változtatni lehessen a kamerából érkező kép felbontását.
2. Folyamatosan bekapcsolva lehessen tartani a felvétel során a kamerához tartozó LED vakut (éjjeli üzemmód).

Ezen funkciókat csak saját kamera kezelő osztály megírásával lehet elérni. Ennek megvalósítását követően a felbontás váltása hardver szinten gyorsan, kevesebb, mint egy másodperc alatt végbemegy, illetve folyamatos üzemben tudjuk működtetni a kamerához tartozó LED vakut.



7. ábra: Dinamikus arckövetés

V. A RENDSZER MŰKÖDÉSI FÁZISAINAK VIZSGÁLATA

A rendszer feladata, hogy megjelenő arcokat „ismert” vagy „ismeretlen” címkékként lásson el. Ehhez azonban először egy felügyelt tanulási fázison kell átesnie, amely során elsajátíthatja a később felismerendő arcot. Ezt az alkalmazás első indításakor teszi meg. Maga a tanulás a később felismerni kívánt jellemzők kivonását, elmentését jelenti. Jelen megvalósításban ezek a leírók az arcról kivont mikrojellemzők: száj, orr, bal szem, jobb szem, szempár. Ez az öt kép kerül elmentésre egy referencia arcról, minden arcrész egy előre megadott egységes méretre skálázva. A rendszer ezen mikroképeket egy napig gyűjti, fontos, hogy ez az idő alatt csak a később felismerni kívánt felhasználni jelenjen a felügyelt területen.

A képkészítések során homályos, elmosódott képek sokszor készülnek. Ennek kiszűrése érdekében egy referencia helyett több kerül elmentésre minden arcrészről. A korrelációt zavarja a fény intenzitásának és irányának változása, ezért ezt kiküszöbölendő, a rendszer egy nap során, több alkalommal is készít referenciákat, időben elszeparálva. Ha egy lokális referencia gyűjtési időben már megvan a kellő számú minta minden mikrojellemzőről, akkor a rendszer tovább folytatja a képek gyűjtését, és az itt kivont jellemzőket azonnal össze is hasonlítja az éppen rögzített referencia jellemzőkkel. Egy személy adott arcrészét sajátjával hasonlítjuk össze. Ezen autokorrelációs értékekből sok készül el. Ezekből átlagot számolva egy dinamikusan meghatározott küszöbszintet

állítunk elő a későbbi felismerés számára, melyet időben hozzárendelünk a referenciákhoz. Egy referenciacsoporthoz és a hozzá rendelt küszöbszinteket láthatjuk a 8. ábrán.



8. ábra: Rögzített referenciacsoporthoz dinamikusan meghatározott küszöbszintekkel.

Amikor a tanulás véget ér, a felügyelő fázis kezd meg munkáját. Ebben a módban, a rendszer folyamatosan, bizonyos időnként referencia betöltést hajt végre. Ez a dinamikus referencia betöltés. Ennek során az elmentett referenciákat végignézve dönt arról, hogy a betöltés idején (valószínűsítőleg) melyek azok, amelyek a legjobb illenek az akkori fényviszonyokhoz (intenzitás, irány).

A felügyelés során folyamatos mozgásérzékelés történik. Ha mozgás jelenik meg a képen, akkor a rendszer egyből emberi alak detektálásra vált át. Ha nem talál embert, akkor bizonyos idő után visszaáll mozgásérzékelésre. Ha talál, akkor a dinamikus arckövető eljárással arc után kezd kutatni. Ha nincs arc, akkor jelszót kér bemondásra, hiszen arc nélkül a felismerő rendszer nem működhet.

Arc detektálása után az arról készült kép elmentése kerül. Adott számú mentett arc után a döntéshozás algoritmusát lát munkához. A referencia és aktuálisan kivont mikroképek normalizált keresztkorrelációja kiszámításra kerül a következő formula alapján:

$$R(x, y) = \frac{\sum_{x', y'} (T(x', y') * I(x + x', y + y'))}{\sqrt{\sum_{x', y'} T(x', y')^2 + \sum_{x', y'} I(x + x', y + y')^2}}$$

ahol T a referencia kép, I pedig az aktuális, összehasonlítandó kép. Esetünkben egy korrelációs értéket kapunk az összehasonlítás eredményeként, mivel a képeink megegyező méretűek. Innen egy lokális döntés:

$$\text{lokális döntés} \begin{cases} \text{igen} & \text{ha } R(T, I) \geq \text{küszöbszint} \\ \text{nem} & \text{ha } R(T, I) < \text{küszöbszint} \\ \text{ismeretlen} & \text{sikertelen jellemzőkinyerés} \end{cases}$$

ahol $R(T, I)$ a referencia és az összehasonlítandó kép keresztkorrelációjának értéke, a küszöbszint pedig a tanulási fázisból dinamikus, autokorrelációkból származtatott érték. Innen egy arcra vonatkoztatott döntés:

$$\text{arc szintű döntés} \begin{cases} \text{extra igen} & \text{ha } i \geq e * n \\ \text{igen} & \text{ha } i > t * n \\ \text{nem} & \text{ha } i < t * n \\ \text{extra nem} & \text{ha } i < (1 - e) * n \\ \text{ismeretlen} & \text{ha } u \geq (1 - e) * n \end{cases}$$

ahol: i azon lokális döntések száma, amelyek eredménye „igen”, e egy magas érték a $[0,1]$ -ből, t alacsony érték $[0,1]$ -ből, $n = rf * nf$, nf a jellemzők száma, rf a jellemzőnként rögzített képek száma, u az „ismeretlen” lokális döntések száma. Jelenlegi megvalósításban: $nf=5$, $rf=5$, $e=0,8$, $t=0,5$. Innen a végleges döntés:

$$\text{végleges döntés} \begin{cases} \text{igen} & \text{ha } e_i \geq w * f \\ \text{nem} & \text{ha } e_n \geq w * f \\ \text{ismeretlen} & \text{ha } e_i = 1 \text{ és } n_i < t * (f - 1) \\ \text{igen} & \text{ha } e_i = 1 \text{ és } n_i \geq t * (f - 1) \\ \text{ismeretlen} & \text{ha } e_n = 1 \text{ és } n_n < t * (f - 1) \\ \text{nem} & \text{ha } e_n = 1 \text{ és } n_n \geq t * (f - 1) \\ \text{igen} & \text{ha } n_i \geq h * f \\ \text{nem} & \text{ha } n_n > (1 - h) * f \\ \text{ismeretlen} & \text{egyébként} \end{cases}$$

ahol: f az arci képek (döntések) száma, e_i az „extra igen”, e_n az „extra nem”, n_i az „igen”, n_n a „nem” arconkénti döntések száma, w egy $[0,1]$ -ből származó kis értékű szám, t ugyanaz, mint korábban, h egy $[0,1]$ -ből származó szám. Jelenlegi megvalósításban $h=t$.

A döntéshozó „nem” kimenete esetén a rendszer egy e-mail illetve MMS üzenet küldését vonja maga után, amely tartalmaz egy képet az ismeretlen arcáról. A riasztást követően ismét mozgásérzékeléssel folytatódik a felügyelet.

A döntéshozó „ismeretlen” kimenete esetén a rendszer hangjelzéssel kéri a megfigyelt területen tartózkodó személyt, hogy mondja be az azonosításhoz szükséges jelszót. Ennek ellenőrzésével teszi „igenné” vagy „nemmé” a bizonytalan kimenetet. Ha a jelszó nem megfelelő, a fentebb leírt riasztás indul el.

A döntéshozó „igen” kimenetére a rendszer nyugtázza a döntést és mozgásérzékeléssel folytatja munkáját. Egy elkészült döntést tartalmazó kép a 9. ábrán tekinthető meg (a képen a bal felső sarokban az éppen aktuális küszöbszintek arcrészek szerint feltüntetve, alatta a végleges döntés, alatta a lokális döntések, arconként oszlopba gyűjtve láthatók).



9. ábra: Döntéshozás eredménye

VI. A RENDSZER TESZTELÉSE

Az elkészült felügyelő rendszer tesztelése több módon és fázisban zajlott. Külön-külön teszteltük az egyes komponensek megbízhatóságát, majd az együttműködésük eredményét. Némely fázist több készüléken is vizsgáltunk. Ezek közt szerepelt: Pulid F11, 2x1GHz elérhető maximális felbontás 1280x720, ár: 20 000Ft (kínai); LG G3, 4x2,5GHz,

használt felbontás 2048x1080, ár: 150 000Ft; Samsung S Advance, 2x1GHz, elérhető maximális felbontás 800x600, ár: 40 000Ft,

Az egyes komponensek vizsgálata során mindegyik gyors működést és nagyon jó eredményeket ért el. Magának az arcfelismerő algoritmusnak a tesztelése volt a másik fő feladat. Az ütemezett referencia készítésnek és betöltésnek a tesztelése sikeres volt. A személyek felismerését több módon is teszteltük. Az első tesztelési forma statikus állapotban, a referencia készítésének helyén történő mérés volt. A statikus állapot a személyre igaz, vagyis nem mozgott a képek készítése során, és a referencia készítés helyén és időpontjában történt a tesztelés. Ez a legjobb eshetőség, hiszen pontosan azokat az anomáliákat szünteti meg, amelyek a korrelációt megzavarják. A tesztek során a felhasználó szerepelt a képeken, tehát a válasz mindig „igen” kellett volna, hogy legyen. Az eredmények az 1-4. táblázatokban figyelhetők meg.

A nehezebb felismerési forma, amikor a személy nem vesz tudomást a készülékről és nem is néz bele, illetve mozgást végez, amely miatt a képek sem élesek. Nyilvánvaló, hogy itt nagyobb szerepet kap a hang alapú ellenőrzés. Maga a tesztelés egy megtanult mozgási sor végrehajtásával történt, hogy az eredmények összehasonlíthatóak legyenek. A 3. táblázatban szereplő mérések esetén a rendszer felhasználója, a megtanult személy felismerése volt a cél. A 4. táblázatban a mozgó, ismeretlen személy alapú mérések összegzése található.

A tesztelés eredményeit röviden összefoglalva elmondható, hogy az eszköz gyorsasága és minősége nem hat ki lényegesen az eredményekre. Az egyértelmű, hogy statikus helyzetben a rendszer megfelelően működik. Az idegen személyek azonosítása kicsivel gyengébb eredményeket hozott, de ha az illető nem ismeri a jelszót, akkor megfelelő a működés, hiszen a hang alapú azonosítás ilyen esetben a főszerep (de pontosan ezért is lett implementálva). Az eredményeket tekintve az elvárt működést tapasztaltuk, az ismert hiányosságok és korlátok jól látszódnak.

Eszköz	Mérések száma	Igen	Nem	Talán (megerősítve hang alapon)
Pulid F11	30	18	4	8 (6)
Samsung S Advance	20	10	3	7 (6)
LG G3	15	8	4	3 (0)

1. táblázat: Statikus, megtanult személy felismerése

Eszköz	Mérések száma	Igen	Nem	Talán (megerősítve hang alapon)
Pulid F11	20	3	10	7 (1)
Samsung S Advance	20	5	11	4 (0)
LG G3	0	0	0	0

2. táblázat: Statikus, ismeretlen személy felismerése.

Eszköz	Mérések száma	Igen	Nem	Talán (megerősítve hang alapon)
Pulid F11	20	5	6	9 (7)
Samsung S Advance	20	7	7	6 (6)
LG G3	10	2	4	4 (3)

3. táblázat: Dinamikus, ismert személy felismerése

Eszköz	Mérések száma	Igen	Nem	Talán (megerősítve hang alapon)
Pulid F11	20	1	3	16 (2)
Samsung S Advance	20	2	4	14 (0)
LG G3	10	2	3	5 (1)

4. táblázat: Dinamikus, nem ismert személy felismerése

VII. ÖSSZEFOGLALÁS

Kutatásunk célja egy olyan mobil alapú, intelligens, alacsony költségű felügyelő rendszer megalkothatóságának a vizsgálata és létrehozása volt, amely képes felismerni a felhasználó arcát. Sikert értünk el egy olyan rendszert megalkotnunk, amely lehetővé teszi arcokról készült képek gyors gyűjtését, ezzel teret adva további, mélyebb kutatásoknak. A jelen megvalósításban szereplő felügyelő rendszer az arcokról kivont mikrojellemzők (száj, orr, bal szem, jobb szem, szempár) képtérbeli összehasonlításával ítéli meg, hogy ismert arc szerepel-e egy adott képen. Habár sikerült így is megfelelő eredményeket elérnünk, nyilvánvaló, hogy ezen jellemzők nagyon érzékenyek a varianciákra. További fejlesztésként elsősorban ennek javítását tűzzük ki célul. A megvalósított algoritmusok közül különösen jól működött a hang alapú jelszóellenőrzés, amely érzékenységet a külső zajokra még javítani kell, de így is 85% feletti pontosságot ért el a tesztelés során.

A rendszer egyszerűen bővíthető, újabb jellemzők vizsgálata, bevétele könnyedén megoldható. A legtöbb eljárás paraméterezhető, tehát további kutatás folyamán is segítséget nyújt a szabad kísérletezésre. Az alkalmazott detektáló eljárások nagyon hatékonyan működnek, ezzel lehetővé téve az arcok gyors rögzítését. A saját kamera kezelő osztály további funkciók kifejlesztését teszi lehetővé, mint például az éjjeli üzemmód, a LED vaku ugyanis szabadon kapcsolható programból. Az alkalmazás Android alapokon íródott, így szinte tetszőleges eszközre változtatás nélkül feltelepíthető és azonnal használható. Az arcfelismerő funkció jelenleg statikus helyzetben lévő személyek esetén megfelelően biztonságosan, még mozgó egyének esetén, érthető korlátok miatt gyengébben szerepel. A tapasztalatok alapján úgy gondoljuk, hogy lehetséges egy olyan rendszer megalkotása, amely közel minden esetben képes egy adott arcra megfelelő döntést hozni. További bővítési és fejlesztési lehetőségek bőven adódnak: pontosított arcfelismerés, több megtanult arcra működő azonosítás, gépi tanulással támogatott döntéshozás, fejlettebb interakciós formák a felhasználóval, hogy csak néhányat említsünk. A rendszer tömegesebb alkalmazása esetén segítene felhasználni az egyébként kihasználatlan okostelefonokat és egyben segítene megővni otthonainkat, értékeinket és támogatná a bűnüldözést.

REFERENCES

1. OpenCV: www.opencv.org
2. N. Dalal and B. Triggs. Histogram of oriented gradients for human detection, Proc. IEEE Int. Conf. Comput. Vis. Pattern Recog., pp.886 -893 Jun. 2005
3. http://docs.opencv.org/modules/gpu/doc/object_detection.html

4. P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In Proc. IEEE Conference on Computer Vision and Pattern Recognition, 2001.
5. http://docs.opencv.org/modules/objdetect/doc/cascade_classification.html
6. Z. Zivkovic. Improved adaptive Gaussian mixture model for background subtraction, Proc. Int. Conf. Pattern Recognit., pp.28 -31 2004
7. <http://docs.opencv.org/java/org/opencv/video/BackgroundSubtractorMOG2.html>