

Hakkel Tamás

2018.04.22.

Gráf-alapú arcfelismerő algoritmus

Bár a gráfelmélet viszonylag fiatalnak mondható, (legalábbis például a geometriához képest, amit már az ókori görögök is műveltek, a maga 1736-ra datálható születésével nehezen szállhat versenybe), ennek ellenére mára már elég rangos helyet vívott ki magának a matematikai tudományágak között. Ennek egyik oka, hogy újszerű megoldást képes nyújtani régi, korábban megoldhatatlannak tűnő problémára, és kifejezetten hatékony a korunkban legnagyobb kihívást jelentő és egyben legnagyobb haszonnal kecsegtető adat-óceán kordában tartásában, feldolgozásában.

Ilyen probléma például a multimédia-adatok feldolgozása, melyből hihetetlen mennyiség áll elő nap mint nap, ám a képek feldolgozása a klasszikus programozói metódus szerint kudarcot vallott. Tekintve, hogy az ember a külvilágot legnagyobb részben a szemén keresztül érzékeli, ezért erős a motiváció, hogy a rendelkezésre álló kép- és videó-anyagot a lehető leghatékonyabb kiaknázására. A képfeldolgozáson belül is kifejezetten érdekes terület az arcfelismerés, több szinten is hely kell állnia a szoftvernek: 1) fel kell ismernie, hogy a képen egy arc van, illetve lokalizálnia az arcot a képen, 2) az arcot össze kell vetnie a rendelkezésre álló, felcímkézett (vagyis névvel ellátott) arcokkal, és 3) a képen látható embert a komoly számításigény ellenére valós időben be kell azonosítania.

Jelenleg az egyetemen hallgatói projekt keretében folyik arcfelismerő rendszer fejlesztése, melynek célja, hogy azokat a hallgatókat és egyetemi munkatársakat, akik ehhez hozzájárulnak, a rendszer (bizonyos betanulási időszak után) a beléptetőkártya használata nélkül is beengedje. A projektben én is részt veszek, ezért felmerült bennem, hogy jelen esszéhez is ebben a témakörben

keresek témát. Nem is kellett sokat keresgélnem, mivel hamar rátaláltam egy 2014-es cikkre, mely azt tűzte kifeladataul, hogy az arcfelismerő rendszert gráfelméleti alapokra helyezze, és így növelje a rendszer hatékonyságát a számítási igény alacsonyan tartása mellett (Kafai, An and Bhanu 2132-2143).

A kutatócsoport az akkor elérhető arcfelismerő rendszereket elemezve a következő általános problémákat sorolták, melyeket az új rendszereknek meg kell oldaniuk: 1) A valós életbeli alkalmazásoknál az arcok mindig más és más pozícióban és orientációban jelennek meg, 2) a megvilágítás folyamatosan változik az arcok rögzítése során, illetve 3) a környezet mellett maga az arc is változik (például különböző érzéseket tükrözve). Az arcfelismerő rendszerek általában geometriai hasonlóságokat keresnek, így jogosan merül fel a kérdés, hogy hogyan lehetne függetleníteni a fentebb sorolt tényezőktől a arcfelismerést. A cikk írói ezért ötvözték a korábbi módszereket gráfelméleti módszerekkel remélve, hogy az arcok közötti hasonlóságok és különbségek jól reprezentálhatók gráfokkal.

A képeket ezért régiókra osztották, melyeket aztán tovább osztottak kisebb cellákra, majd a cellákból feature-vektorokat állítottak elő. Ezt megismételték a többi felcímkezett képre is, majd az így előállt vektor-tömeget emberek szerint csoportosították. Az egy emberhez tartozó vektorok együtt egy "referencia-arcot" alkottak, mely az ún. "referencia-arcok gráfjának" voltak a csúcspontjai. A gráfot teljes, súlyozott gráfként definiálták, melyben az élek súlyát az él által összekötött két vektor-halmaz hasonlósága adta. Két vektor-halmaz közötti hasonlósági metrikát pedig úgy definiálták, hogy a cellák feature-vektorainak DCT (Discrete Cosine Transform) hashek egyezését számlálták meg régióként, és a legnagyobb egyezést mutató régiók közti hash-egyezések összege adta meg a mérőszámot. Következő lépésben az így létrehozott gráfon belül aztán kiszámolták minden egyes csúcsra a centralitást. Háromféle centralitást is használtak:

A kimenő élek súlyának összegéből adódott egy, és ezen kívül betweenness és closeness centralitási értékeket számoltak.

Ezzel le is zárult az előkészítő szakasz, innentől a gráf használható arcfelismerésre. Ennek kivitelezéséhez először a beazonosítandó képalapján a "referencia-arcokhoz" vagyis a gráf csúcspontjaihoz egyenként számoltak egy-egy "bázis leíró" ("basis descriptor"). Ez a vektor annyi elemű volt, mint a referencia-arcok száma (nevezzük ezt a számot N -nek), és minden egyes eleme a beazonosítandó arc (korábban tárgyalt hasonlósági metrika szerint számolt) távolságát tárolta az adott pozíciónak megfelelő referencia-arctól. A referencia-arcok centralitás értékeit ezután egy $3 \times N$ -es mátrixba rendezték, melynek három sora a három különféle centralitás típusnak felel meg, és ennek minden sorát megszorozták elemenként a "base descriptor"-ral. Az így kapott vektorból pedig már könnyen eldönthető, hogy melyik "reference face"-nél szerepelnek a legnagyobb értékek, vagyis a beazonosítandó képen ki szerepel az arcfelismerő szerint.

A módszer nagy előnye, melyet a cikk végi kiértékelések is alátámasztottak, hogy cellákra osztás miatt a rendszer robosztusnak bizonyult nagy mértékben különböző arcbeállítások mellett is, és emellett a DCT hash-nek köszönhetően $O(1)$ idő alatt lehet összehasonlítani referencia-arcokkal (szemben a korábbi módszerekkel, ahol az egy emberhez tartozó képek növelésével lineárisan nőtt a futási idő is). A kiértékeléseket több adatbázison elvégezték és számos state-of-the-art algoritmussal összehasonlították. A kiértékelés megmutatta, hogy egyfelől (nem meglepő módon) a képek számával nő a pontosság is, ami 92% körül szaturálódik, illetve a többi algoritmussal összevetve kiderült azok bonyolult preprocessz lépésekkel is lemaradtak a módszer eredményei mögött.

4 év azonban már jelentős időnek számít az informatikának ezen ágában, így már most érződik, hogy a korábbi módszereket szép lassan kiszorítják a mélytanuló algoritmusos

módszerek. Jelen arcfelismerős projektünkben is neurális hálózatot használunk, melynek a teljesítménye messze meghaladja a cikkben leírt módszerét. A két módszer ötvözése viszont lehetséges, mivel a neurális hálózatunk is feature-vektorokat gyárt a képekből, melyek a jelen esszében bemutatott algoritmus bemenetét képezik, így ez egy lehetséges továbbfejlesztési irány lehet a továbbiakban.

Works Cited

Kafai, Mehran, Le An, and Bir Bhanu. "Reference Face Graph For Face Recognition." *IEEE Transactions on Information Forensics and Security*. 9.12 (2014): 2132-2143. Web. 23 Apr. 2018.