

СТАНОВИЩЕ

от проф. д-мн Огнян Иванов Кунчев,
член на Научно жури
относно дисертационния труд на тема „ВИДЕОСТАБИЛИЗАЦИЯ И 3D
РАЗПОЗНАВАНЕ В РЕАЛНО ВРЕМЕ“
на
Атанас Филипов Николов
за придобиване на образователната и научна степен „доктор“
към ИИКТ, БАН– София

1. Биографични данни:

Кандидатът е завършил специалността Компютърни системи и технологии в ТУ-София, с магистърска степен придобита през 2010 г. Магистърската му теза е в областта на Обработка на изображения и Компютърно виждане.

Настоящият дисертационен труд е естествено продължение на темата на тази работа.

Дисертантът е бил редовен докторант към ИИКТ-БАН в периода 2011-2014, като темата на неговата докторантура в ИИКТ е била Разпознаване на образите и Биометрика – Разпознаване на уши и Видеостабилизация. Всички тези биографични данни показват естествената последователност, в която докторантът е придобивал знания и умения, и в които е била извършена работата по дисертацията. Съществена роля за качеството на работата е била средата в ТУ-София и ИИКТ, създадена от едни от най-изтъкнатите специалисти у нас в областта на Обработка на изображения.

2. Представяне на тематика и анализ на постигнатите резултати.

Темата на дисертацията е Видеостабилизация и тримерно (3D) Разпознаване в реално време. Преди да разгледаме основните достижения на дисертанта, нека споменем научната актуалност на разглежданите проблеми.

Дисертационният труд е посветен на едновременното решение на два проблема – от една страна това е проблемът за двумерна (2D) видеостабилизация в реално време, а от друга страна това е проблемът за (3D) тримерно разпознаване на личността по формата на ушите.

Относно първата област, трябва да подчертаем, че методите за видеостабилизация са една много важна съвременна област на практическото приложение на Обработка на изображенията, която цели отстраняването на естествените шумове/трептения, предизвикани от качествата на нестабилно заснети видеокадри „в точка“ или „по траектория“. Технически погледнато, стабилизацията „в точка“ означава привеждането на всички видео кадри към един референтен кадър, т.е. да се елиминират (ротационните) трептения на камерата „в точка“. В резултат на това се получава видео, което е заснето от една (виртуална) статична камера. Целта на „Стабилизацията по траектория“ на едно видео е да се получи видео, при което „виртуалната“ камера следва изгладената оригинална траектория на оператора. Решаването на тези две задачи биха били от голямо значение например в киноиндустрията, където за стабилното заснемане се използва голямо разнообразие от сложно и много скъпо оборудване: стативи; камери, движещи се върху релси; операторски колички; преноси-ми стабилизиращи механизми, закрепени към тялото и др. Поради високата цена на оборудването, за любителски цели широко се използват софтуерни методи за постигане на гладкост на движението на камерата.

Известните днес подходи, които решават задачата за софтуерната 2D стабилизация и дори 3D стабилизация, имат сериозен проблем с изчислителното време. Най-известните реализации са тези на *Google YouTube Stabilizer* и *Warp Stabilizer* на *Adobe After Effects*, но те не могат да достигнат достатъчно бързодействие на PC от среден клас. За компенсация на недостатъчното бързодействието на алгоритмите, при преносимите устройства се разчита на вградени в хардуера инерциални сензори (жироскопи и/или акселерометри) комбинирани с прецизни и бързи софтуерни подходи. Поради това е необходимо създаването на нови бързи методи за видеостабилизация, и **настоящата дисертация дава своя принос в това направление.**

В Глава 2 на дисертацията се разглеждат софтуерни методи за 2D видеостабилизация както по гладка траектория, така и в „точка“. Тези методи са подходящи за вграждане в мобилни телефони, тъй като имат достатъчно бързодействие за работа в реално време и произвеждат задоволително качество на образа. Едно интересно приложение е получено в областта на откриване на рекламни видеоклипове. Предложен е бърз, прецизен и относително прост алгоритъм, работещ в реално време, който дава възможност да се локализира предварително известни рекламни видеоклипове и/или произволни части от тях в базов TV видеопоток. (Това ще реши задачата за следенето на излъчването на реклами според условията на договора с дадената телевизия, което в момента се осъществява от хора (оператори), които наблюдават различни TV канали в съответните предварително фиксирани времеви интервали за излъчване на реклами.)

Основно научно постижение на дисертанта в тази част на дисертацията е подобрение на известния SAD подход. Подобрението се състои в това, че вместо класическите проекции, се провеждат изчисления върху нормализираните по средна стойност проекции, при това се допуска средната стойност да е „плаваща“. Проведени са експерименти върху 5 видеоклипа (от различни шотове), с размер 640x480 и с различна степен на трептене, като са тествани различни методи за видеостабилизация.

В тази част на дисертацията дисертантът е показал умение да работи с основните технически средства на видеостабилизацията, като например твърдотелен (CCD) модел на трансформация между кадрите и „rolling shutter“ (CMOS) модел. Проведените експерименти върху тестови видео-клипове са обещаващи за бъдещото развитие на метода, като се очаква той да се комбинира с известните методи за видеостабилизация, използващи инерциални сензори, за постигане на прецизна видеостабилизация.

Втората област на дисертацията е посветена на Методи за разпознаване на индивиди по формата на ушите. Това е една много силно акцентирана област напоследък, тъй като ред изследвания показваха, че формата на ушите е уникална за индивида, и това ги прави отличителна характеристика за биометрично разпознаване на хора. От големия брой изследователски статии по темата от последните десетилетия може да се заключи, че точността, постигната чрез системите за разпознаване по уши, е съизмерима с тази на разпознаването по лица. Например, според едно изследване през 2005 г. в Университета на Southampton, UK, относно индекса на отличимост „*decidability index*“, който измерва подобие на ушите като показател за степента на уникалност на човешките уши, е установено, че този индекс е с един порядък по-голям за уши, отколкото за физиономии. Известно е също, че формата на ухото еволюира по време на ембрионалния растеж от шест развиващи се базови възела. Освен това, структурата на ушите не се променя значително с възрастта.

В Глава 3 на дисертацията е описана разработената многомоделна БД (3DEarDB), съдържаща различни моделни представяния на десните уши на около 100 индивиди. Разгледан е известният метод на Разширено Гаусово изображение EGI (Extended Gaussian Image) за представяне на 3D обекти. РГИ представянето е предложено от В.К.Р. Horn през 1984 г., за еднозначно представяне на **изпъкнала** 3D повърхност.

Тъй като има най-разнообразни формати за представяне на формата на човешкото ухо, целта на Базата данни 3DEarDB е да осигури съвместимост между различните изходни формати за представянето на ухото на един и същ човек. Тези формати включват: (i) сурови (необработени) 3D модели на уши; (ii) обработени 3D модели; (iii) Kinect дълбочинни

изображения "depth maps"; (iv) 2D видео клипове; (v) генерирани структури от 2D интензитетни проекции на уши; (vi) генерирани структури от 2D дълбочинни изображения. Създадената база данни е един **съществен принос на дисертанта**, който може да бъде по-нататък използван от други изследователи в областта на разпознаване по ушите, за тестване и сравнение на точността на алгоритмите върху различни възможни сценарии.

В тази част на изследването **дисертантът е показал умения** да работи с различните формати на представяне на човешкото ухо, в частност с РГИ представянето, и да внася нови елементи в тази техника. Например, експериментирана е комбинация от възможности за разпознаването на ушите по метода на най-близкия съсед, симулиращи различна резолюция на 3D скенера, както и различна сложност на РГИ представянето. Направено е интересното експериментално откритие, че разликите между формите на ушите на два (индивида) са достатъчно големи, така че нееднозначността на РГИ представянето от изпъкналости и вдлъбнатости на формата на ушите не влияе върху разпознаемостта.

Като цяло, трябва да се отбележи, че в работата са проявени уменията на дисертанта в областта на Линеината алгебра и Аналитичната геометрия, Реалния анализ, Статистиката, Обработка на изображения, Разпознаване на образи, и Компютърно зрение. Това показва, че образователната компонента на степента доктор е постигната.

Проектна и други дейности на дисертанта

Дисертантът е взел участие в 5 научно-изследователски проекта на ИИКТ-БАН, а също и в 3 други проекта, където е извършена и апробация на резултатите.

Анализът на представените материали говори за това, че кандидатът има високо ниво на теоретични познания в редица области на Приложната математика и ИКТ. Не по-малко важни са и неговите умения и ориентираност за прилагане на тези познания към конкретни практически проблеми в областта на Обработка на изображения.

3. Забележки:

В автореферата не са коректно дадени номерата на цитираната литература.

4. Наукометрични данни

Представени са 5 излезли от печат и една приета за публикуване статия; всичките са съвместно с други съавтори. Забелязани са и 3 цитата на публикациите.

ЗАКЛЮЧЕНИЕ

Анализът на приложените материали характеризират дисертанта като млад и перспективен учен с добра професионална подготовка и доказани изследователски възможности в областта на Приложната математика, ИКТ и Обработка на изображенията.

На базата на всичко гореизложено, препоръчам на Уважаемото научно жури да присвои образователната и научна степен „доктор“ по Докторска програма: „Компютърни системи, комплекси и мрежи“, в Професионално направление: 5.3 „Комуникационна и компютърна техника“

Дата: 18.05.2016 г.

ЧЛЕН НА ЖУРИТО: