

Зубенко І.М.¹, Корначевський Я.І.¹, Макеєнок О.М.²

¹ННК “Інститут прикладного системного аналізу” НТУУ “КПІ”, Київ, Україна; ²Міжнародний науково-навчальний центр інформаційних технологій та систем НАНУ та МОНМСУ, Київ, Україна

Сучасні інформаційні технології програмної обробки відеопослідовностей

Сьогодні програмна обробка відео може здійснюватися двома шляхами: за допомогою відеоредакторів, і за допомогою бібліотек програм обробки відео. Значна частина відеоредакторів дозволяє використовувати додаткові програми – плагіни, які розширюють можливості редакторів. Часто користувач може розробляти власні плагіни, які реалізують потрібні саме для нього алгоритми. Бібліотеки обробки відео надають користувачу набір стандартних алгоритмів, який користувач може при необхідності розширити, створюючи нові власні функції (як за допомогою вже існуючих функцій, так і за допомогою звичних йому мов програмування, наприклад С, або С++).

Нижче дається аналіз основних відеоредакторів та бібліотек обробки відео.

В дослідженні були розглянуті наступні редактори відео: OpenShot, VirtualDub + Avisynth, Cyberlink PowerDirector, PiTiVi, Cinelerra, Kdenlive.

OpenShot – це редактор, заснований на python + FFmpeg, і, таким чином, є крос-платформним і його просто модифікувати і розширювати. Він надає доступ до різних даних кадру і може бути використаний для широко класу задач обробки відео.

VirtualDub + Avisynth (скриптовий інтерпретатор) – це пара, яка містить простий редактор відео, а також кадровий сервер. Вони дозволяють використовувати велику кількість існуючих фільтрів. Продуктивність цих програм дозволяє робити це майже в реальному часі. Більшість фільтрів добре задокументована, вони мають відкриті програмні коди і можуть бути просто пристосовані до потреб користувачів.

Cyberlink PowerDirector – це простий редактор для новачків, який дозволяє обробляти відео за допомогою накладання шарів, які реалізують найбільш поширені алгоритми, але він не дозволяє підвищити його можливості за допомогою додавання власних програмних модулів користувачів.

PiTiVi та Cinelerra – це альтернатива парі VirtualDub+Avisynth, вони обидва мають багатий набір фільтрів та можливостей, але основна проблема полягає в складності їх використання і невеликій кількості користувачів.

Kdenlive являє собою нелінійний відеоредактор, який охоплює застосування від початкового до напівпрофесійного рівня. Він підтримує кілька відеотреків. Однак набір фільтрів (ефектів) не такий широкий, як у конкурентів. Реалізація ефектів заснована на API плагінах. Документація є недостатньо детальною.

Висновок: до дослідницької обробки відео доцільніше використовувати VirtualDub + Avisynth, у зв'язку з наявністю великого набору фільтрів, швидкою роботою та простотою програмування.

В дослідженні були також розглянуті наступні бібліотеки для роботи з відео:

OpenCV (<http://sourceforge.net/projects/opencvlibrary/>) – бібліотека з відкритим програмним кодом (ліцензія BSD), призначена для обробки зображень і потокового відео. Написана на мовах С і С++, що дозволяє використовувати її в портативних пристроях. Ця бібліотека офіційно підтримується на операційних системах Windows, Linux, Mac OS X, Android. Кодування і декодування відеопотоків бібліотека OpenCV виконує за допомогою бібліотеки FFmpeg. Бібліотека OpenCV призначена для створення програм, які працюють в режимі реального часу, надає користувачу прямий доступ до відео кадрів та дозволяє отримувати потік відео прямо з камери. Бібліотека розділена на 12 компактних модулів, які містять понад 500 функцій.

core – структури даних і функції, які реалізують зберігання зображень, накладення масок,

складання зображень, обчислення ШПФ, налаштування яскравості і контрасту, малювання.

imgproc – фільтрація зображень, геометричні перетворення, зміна колірних просторів, обчислення гістограм, поліпшення якості зображень, робота з порога, перетворення Собеля, Лапласа, морфологічні та афінні, визначення країв, зіставлення зі зразком.

highgui – графічний інтерфейс та функції вводу-виводу відео з файлів та камер.

video – функції виявлення і відстеження рухомих об'єктів, відокремлення їх від фону, передбачення наступної позиції об'єкта.

calib3d – калібрування відеокамери і 3D-реконструкції – багатовимірної реконструкції об'єктів на основі їх двовимірних зображень.

features2d – відображення зображення (mapping), пошук кутів методами Харріса і Ши-Томасі або методами користувача, пошук заданих об'єктів, пошук плоских поверхонь.

objdetect – виділення окремих об'єктів у відеопотоці і пошук зазвичай використовуваних об'єктів, таких як люди, автомобілі, будинки і дерева.

ml – машинне навчання – функції, які використовуються в робототехнічних системах.

flann - функції кластеризації та пошуку об'єктів в багатовимірних просторах.

gpu - оптимізовані функції для графічних процесорів, наприклад, NVIDIA CUDA.

Бібліотека добре документована. Доступні книги з її використання та приклади програмного коду.

Aforge.NET (<http://www.aforgenet.com>). Ліцензія: LGPL. Бібліотека реалізована на мові C# в середовищі Microsoft NET і призначена для розв'язання різноманітних задач, у тому числі обробки зображень і роботи з потоковим відео. У бібліотеці є класи, спеціально призначені для виявлення рухомих об'єктів та їх відстеження. Якість виявлення і відстеження об'єктів добра.

VXL (<http://vxl.sourceforge.net>) – це набір з 12 взаємопов'язаних бібліотек, призначених для обробки потокового відео, наукових досліджень та створення реальних систем в галузі машинного зору. Бібліотека реалізована в ANSI C++ як крос-платформенне рішення з відкритим програмним кодом. Проект в даний час активно розвивається. Великий розмір коду бібліотеки ускладнює його використання в портативних пристроях.

NuPic (<http://numenta.com>) – одна із сучасних спроб побудувати діючу модель людського мозку. Використання такої моделі для вирішення конкретної проблеми вимагає її обов'язкового попереднього навчання. Автори докладно описують теоретичні основи створення моделі і дають керівний документ для її практичної реалізації у вигляді програмного продукту. Як приклад використання бібліотеки автори наводять виявлення рухомих об'єктів у відеопотоці.

FFMPEG (<http://ffmpeg.mplayerhq.hu>) – реалізована на мові C та складається з 4 основних частин: libavcodec – бібліотека відео та аудіо кодеків, libavformat – бібліотека процедур читання та запису аудіо та відеофайлів практично всіх сучасних форматів. libavfilter – бібліотека фільтрів обробки відео- та аудіопотоків. Містить 40 відео фільтрів та дозволяє додавати фільтри, розроблені користувачем, ffmpeg – утиліта командного рядка, яка дозволяє використовувати наведені вище можливості бібліотеки для обробки аудіо і відео без необхідності написання нового коду.

Здійснене дослідження показує, що в даний час для дослідницької обробки відео найкраще підходить бібліотека OpenCV. Ця бібліотека є надійним рішенням, воно підтримане авторитетною командою розробників, забезпечує найбільші можливості обробки відео потоків, є крос-платформеним рішенням, доступна в програмних кодах під ліцензією BSD і має кращу, порівняно з іншими, документацію.