

UNIVERZA V LJUBLJANI
FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Jernej Ulčakar

Uporaba globinske kamere za učenje igranja klavirja

DIPLOMSKO DELO

UNIVERZITETNI ŠTUDIJSKI PROGRAM
PRVE STOPNJE
RAČUNALNIŠTVO IN INFORMATIKA

MENTOR: doc. dr. Luka Čehovin Zajc

Ljubljana, 2022

To delo je ponujeno pod licenco *Creative Commons Priznanje avtorstva-Deljenje pod enakimi pogoji 2.5 Slovenija* (ali novejšo različico). To pomeni, da se tako besedilo, slike, grafi in druge sestavine dela kot tudi rezultati diplomskega dela lahko prosto distribuirajo, reproducirajo, uporabljajo, priobčujejo javnosti in predelujejo, pod pogojem, da se jasno in vidno navede avtorja in naslov tega dela in da se v primeru spremembe, preoblikovanja ali uporabe tega dela v svojem delu, lahko distribuira predelava le pod licenco, ki je enaka tej. Podrobnosti licence so dostopne na spletni strani creativecommons.si ali na Inštitutu za intelektualno lastnino, Streliška 1, 1000 Ljubljana.



Izvorna koda diplomskega dela, njeni rezultati in v ta namen razvita programska oprema je ponujena pod licenco GNU General Public License, različica 3 (ali novejša). To pomeni, da se lahko prosto distribuira in/ali predeluje pod njenimi pogoji. Podrobnosti licence so dostopne na spletni strani <http://www.gnu.org/licenses/>.

Besedilo je oblikovano z urejevalnikom besedil L^AT_EX.

Kandidat: Jernej Ulčakar

Naslov: Uporaba globinske kamere za učenje igranja klavirja

Vrsta naloge: Diplomaska naloga na univerzitetnem programu prve stopnje
Računalništvo in informatika

Mentor: doc. dr. Luka Čehovin Zajc

Opis:

Cilj naloge je izdelava interaktivnega sistema, ki sestoji iz projektorja in globinske/navadne kamere. Sistem zazna klaviaturo ter nanjo projecira informacije, ki omogočajo lažje učenje igranja skladb (npr. projecira pritiske tipk v časovnem zaporedju). V nadaljevanju se lahko sistem nadgradi tudi z detekcijo rok.

Title: Using depth camera to learn piano

Description:

The aim of the task is to build an interactive system that includes a projector and a depth or a normal camera. The system detects the piano keyboard and projects information onto it, which enable easier learning of playing music (for example press of a button in a time order). In the future the system can be upgraded with hand detection.

Kazalo

Povzetek

Abstract

| | | |
|----------|--------------------------------|-----------|
| 1 | Uvod | 1 |
| 2 | Motiv aplikacije | 3 |
| 3 | Algoritmi za detekcijo | 5 |
| 3.1 | Hough Transform | 5 |
| 3.2 | SIFT | 6 |
| 4 | Celotna integracija | 7 |
| 4.1 | Problemi | 7 |
| 5 | Rezultati | 9 |
| 5.1 | Detekcija klaviature | 9 |
| 5.2 | Detekcija rok | 9 |
| 5.3 | Notni stolpci | 9 |
| 6 | Zaključek | 11 |
| 6.1 | Nadgradnja | 11 |

Seznam uporabljenih kratic

| kratica | angleško | slovensko |
|-------------|-----------------------------------|--|
| SIFT | scale invariant feature transform | metoda za detekcijo značilnosti na sliki |

Povzetek

Naslov: Uporaba globinske kamere za učenje igranja klavirja

Avtor: Jernej Ulčakar

Diplomska naloga je namenjena za igralce klavirja, ki si upajo in želijo poskusiti nov način učenja skladb. Prav tako je veliko bolj zabavna varianta igranja, če je oseba vešča o arkadnih igrah. Izdelek naj bi uporabljal algoritme za razpoznavo tipk in celotne klaviature z detekcijo robov ter za detekcijo prstov s pomočjo blob detection algoritmov. Oseba, ki preizkuša izdelek mora biti opremljen s kamero in monitorjem, na kateremu bo prikaz celotne aplikacije.

Ključne besede: računalnik, kamera, klavir.

Abstract

Title: Diploma thesis template

Author: Jernej Ulčakar

Thesis is meant for piano users, who wish to try out new ways of learning piano music. It is a more fun way of playing if the person likes arcade video games. The final product should use algorithms to detect hands and piano keyboard using edge detection and blob detection algorithms. The user should have a standard camera and a monitor or a projector to project the whole application .

Keywords: computer, camera, piano.

Poglavje 1

Uvod

Namen diplomske naloge je ustvariti neko delujočo aplikacijo, ki pripomore uporabniku, da se nauči igrati klavir brez poznavanja not in glasbene teorije. Končni uporabnik je lahko kdorkoli, ki ima doma ali nekje drugje dostop do klavirja. Možno je tudi preko naprintane klaviature, ki naj bi ustrezala standardnem formatu klavirja.

Za prepoznavanje klaviature bom uporabil algoritme za detekcijo robov ter nato na robusten način odrezal kos slike, ki nosi klaviaturo. Potlej se pripravi scena, ki zavzema približno četrtno zaslona po krajši dolžini zaslona skozi celotno dolžino klaviatura, ostali zaslon pa stolpci, ki padajo proti klaviaturi, in mogoče kakšni dodatni gumbi za opcije in podobno.

V prihodnjih poglavjih je podrobnejša razlaga o posameznih delih razvoja te aplikacije. V 2. poglavju povem o motivu projekta. V 3. poglavju povem o uporabi algoritmov. V 4. poglavju povem o celotni integraciji, torej celotna aplikacija. V 5. poglavju bom povedal še o rezultatih testiranja, ki jih trenutno še ni. Sledil bo samo še zaključek 6.

Poglavje 2

Motiv aplikacije

Temo sem si izbral, ker sem že od otroštva igral klavir in hkrati študiral na fakulteti umetno inteligenco. Rad bi ustvaril lahek pripomoček za igranje klavirja brez poznavanja not in drugih pravil na notnem črtovju.

Rezultati diplomske naloge bodo koristili vsem, ki bi si želeli igrati klavir brez not in na malo drugačen način. Z uporabo umetne inteligence bomo opazovali klaviaturo in na steno projecirali katero tipko naj uporabnik pritisne ob določenem času. Celoten izdelek spominja na arkadne igre, saj mora uporabnik upoštevati tempo glasbe in opazovati projekcijo, kar tudi vpliva na motivacijo izdelave aplikacije.

Poglavje 3

Algoritmi za detekcijo

3.1 Hough Transform

Za detekcijo klaviature se bo uporabil Hough transform algoritem, ki bo vzel tiste črte, ki so najbolj “vodoravne” zaslonu oz. tisti, ki so najboljši kandidati. [hough]

Algoritem 3.1 *Hough Transform*

Hough transform je vrsta tehnike

feature extraction, ki se uporablja v analizi slik in računalniškemu vidu ter v procesiranju slik. Deluje na principu volitev, ki se izvajajo v parametričnem prostoru, kjer so kandidatni objekti pridobljeni kot lokalni maksimum v tako rekočem “Accumulator” prostoru, ki je eksplicitno ustvarjen pri Hough transform algoritmu, da se lahko izvede in izračuna.

3.2 SIFT

Za detekcijo rok se bo uporabil SIFT algoritem s pomočjo ključnih točk.

Algoritem 3.2 *SIFT*

SIFT oz. *Scale-invariant feature transform* je “feature extraction” algoritem, ki se uporablja za računalniški vid za detekcijo in opis lokalnih značilnosti znotraj slik.

SIFT-ove ključne točke objektov oz. “keypoints” so najprej vzete iz množice referenčnih slik in shranjene v neko podatkovno bazo. Objekt v novi sliki je prepoznan s primerjanjem ključnih točk znotraj podatkovne baze, kjer se nato vzame najboljši kandidat z uporabo evklidske razdalje na njegove značilnostne vektorje. Iz celotnega nabora zadetkov so opredeljeni podnabori ključnih točk, ki se strinjajo glede predmeta in njegove lokacije, obsega in usmeritve na novi sliki, da se filtrirajo dobra ujemanja. Določitev doslednih grozdov se hitro izvede z uporabo učinkovite izvedbe zgoščene tabele generalizirane Hough-ove transformacije [3.1]. Vsaka skupina treh ali več lastnosti, ki se strinjajo glede predmeta in njegovega položaja, je nato predmet modela nadaljnjega podrobnega preverjanja, nato pa se odstopanja zavržejo. Končno se izračuna verjetnost, da določen nabor lastnosti kaže na prisotnost predmeta, glede na natančnost prileganja in število verjetnih napačnih ujemanj. Ujemanja predmetov, ki so prestala vse te teste, lahko z veliko zaupanjem prepoznamo kot pravilna.

Poglavje 4

Celotna integracija

Ko so deli naloge napisane v celoti, se vse skupaj združi in prilagodi. Na začetku aplikacije je zahtevano polje za IP naslov kamere in njen naslov vrat (port). Potlej se ob ustrezni povezavi preklopi na sliko kamere, kjer bo zahtevano, da se kamera pravilno postavi, da lahko algoritmi za zaznavanje klaviature delujejo zaželeno. Ko je vse urejeno, se prikaže gumb ali kaj podobnega za pričetek igranja skladbe ali snemanje. Tu se bodo na zaslonu prikazali razni stolpci, ki gladko tečejo proti ustrezni tipki na klaviaturi.

4.1 Problemi

Pri celotni aplikaciji obstajajo različni problemi, ki niso tako enostavno rešljivi, saj so odvisni od računalnikov in kamer ter stabilnostjo povezave.

4.1.1 Latenca

Problem je latenca kamere in latenca povezave ter hitrostjo procesorja zaradi algoritmov za zaznavanje. Celotna latenca traja nekje med 50ms in 200ms. Ta je zelo opazna med igranjem klavirja, če gledamo naše prste. Vendar zadeva ni tako težavna, če se prilagodimo tako, da ignoriramo naše prste na klaviaturi in opazujemo zgolj stolpce ter pripadajoče tipke.

4.1.2 Postavitev

Do težav pride, ko skušamo celoten sistem vzpostaviti tako, da je kamera lepo postavljena točno nad klavirjem, ki gleda v klaviaturo. Če nimamo pri sebi pripomočkov za držo kamere, lahko postane postavljanje celotnega sistema zelo zahtevno ali pa nemogoče.

4.1.3 Projektor

Če oseba ne želi uporabljati zaslona, lahko to reši z uporabo projektorja, ki naj bi kazal na steno ali na prostor za note. Tu je kvaliteta slike odvisna ne le od projektorja, temveč tudi od stene na katero bo projeciral.

Poglavje 5

Rezultati

Rezultati testiranja bodo tukaj še zapisani. Trenutna pričakovanja niso dobra, saj med ustvarjanjem in razvijanjem aplikacije se rezultati spreminjajo.

5.1 Detekcija klaviature

Tukaj bodo zapisani rezultati detekcije klaviature ob postavitvi kamere.

5.2 Detekcija rok

Tukaj bodo zapisani rezultati detekcije rok ob igranju klavirja

5.3 Notni stolpci

Tukaj bodo zapisani rezultati notnih stolpcev, ki se premikajo proti tipkam na klaviaturi.

Poglavje 6

Zaključek

Celoten projekt je zelo zanimiv za implementirati in za testirati, saj lahko poleg testiranja tudi igraš klavir in se s tem sproščaš. Največ časa bo trajala implementacija celotne zaznave, nato pa implementacija skladb.

6.1 Nadgradnja

Nadgradnje sistema so lahko drugačni in bolj robustni algoritmi, ki naj bi zanemarili težavnost pri postavitvi kamere (npr. kamera s strani). Druge nadgradnje so zmanjševanje latence s hitrejšimi algoritmi.