

Porównanie metod SIFT, SURF, BRIEF i ORB w celu rozpoznawania gitar elektrycznych i akustycznych przy różnych warunkach otoczenia

Patryk Gliszczyński, Katarzyna Boczek

Politechnika Poznańska, 2015

Praca ta jest poświęcona tematowi przetwarzania obrazów. Jest to bardzo interesujący i ciągle rozwijający się dział współczesnej informatyki. Dzięki technologii przetwarzania obrazów jesteśmy w stanie uczyć komputer w jaki sposób ma postrzegać rzeczywistość, docelowo dążąc aby była ona postrzegana tak jak widzimy ją my, ludzie. Możemy znajdować dowolne obiekty na otrzymanych obrazach, odczytywać dane z otoczenia i przetwarzać je w dowolny sposób, na przykład tłumacząc widziany tekst na dowolny z języków, odczytywać zapisy muzyczne i automatycznie je odtwarzać, rozpoznawać przeszkody w inteligentnych systemach prowadzenia pojazdów, czy rozpoznawać ludzkie twarze i ich emocje. Zagadnienie to jest więc bardzo otwarte, i nasza praca nie odniesie się w pełni do możliwości tego tematu.

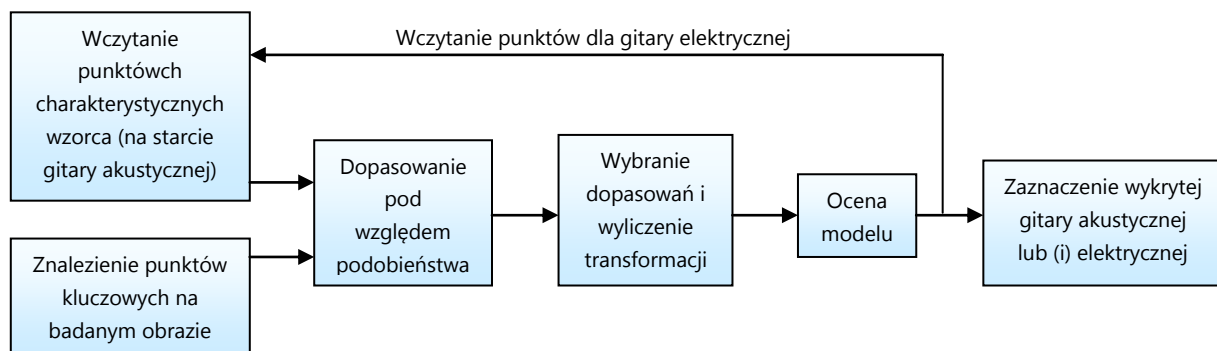
WPROWADZENIE

W naszej pracy skupiliśmy się na rozpoznawaniu dwóch typów gitar: akustycznych i elektrycznych. Na samym początku mieliśmy parę różnych pomysłów jak podejść do tego zagadnienia, np. poprzez operacje progowania, domknięcia i erozji rozpoznawać kontury gitar, i w zależności od rodzaju gitary kontury również były by inne. Uznaliśmy natomiast, że rozwiązanie to nie zda egzaminu, gdy gitary będą sfotografowane pod innym kątem, ich perspektywa się zmieni, lub będą w innej odległości na obrazie (brak skalowalności). Dlatego kolejnym naszym pomysłem było wykrywanie okręgu, który tworzy otwór w pudle rezonansowym gitary akustycznej, którego nie ma w gitarze elektrycznej. Jednak uznaliśmy ten pomysł za mało ciekawy i ostatecznie zdecydowaliśmy się wykrywać punkty charakterystyczne na zdjęciu pokazowym gitary, a następnie poszukiwać tych punktów na analizowanym zdjęciu różnymi metodami. Podczas przeprowadzania testów skorzystaliśmy z czterech dobrze znanych algorytmów wykrywania punktów charakterystycznych z biblioteki OpenCV: SIFT, SURF, BRIEF i ORB. Każda z tych metod jako dane wejściowe przyjmuje zdjęcie, na którym poszukiwać będziemy dany obiekt, oraz zdjęcie wzorcowe (czyli takie, do którego będzie się odnosić w procesie poszukiwania odpowiednich punktów charakterystycznych na obrazie).

Pierwszą metodą, którą wykorzystaliśmy jest **SIFT** (ang. Scale-Invariant Feature Transform), która bazuje na wyodrębnianiu punktów charakterystycznych, które są niezależne od skali, oraz orientacji obrazu. W tym celu stosuje się filtr Gaussa o różnych parametrach. Dla wyznaczonych punktów oblicza się wektor cech będący rodzajem histogramu gradientów liczonych w ich otoczeniu. Cechy te są w dużym stopniu niewrażliwe na zmianę oświetlenia, obrót czy skalę. Wyznaczony wektor jest pewnym

opisem obrazu, dlatego porównując wektory dwóch obrazów możliwe jest znalezienie ich wspólnych cech.

Drugą metodą jest metoda **SURF** (ang. Speeded Up Robust Features), która jest szybszym odpowiednikiem metody SIFT. Podstawową różnicą jest sposób wyznaczania punktów charakterystycznych. W metodzie SURF wykorzystywana jest macierz Hessego, która jest mniej czasochłonna niż wyznaczanie przybliżeń Gaussa. Metoda ta, podobnie jak metoda SIFT, sprawdza się dla obiektów, które mają dużo punktów charakterystycznych. Działa jednak gorzej przy słabszym oświetleniu.



Rysunek 1 Ogólny schemat blokowy przetwarzania w tym problemie

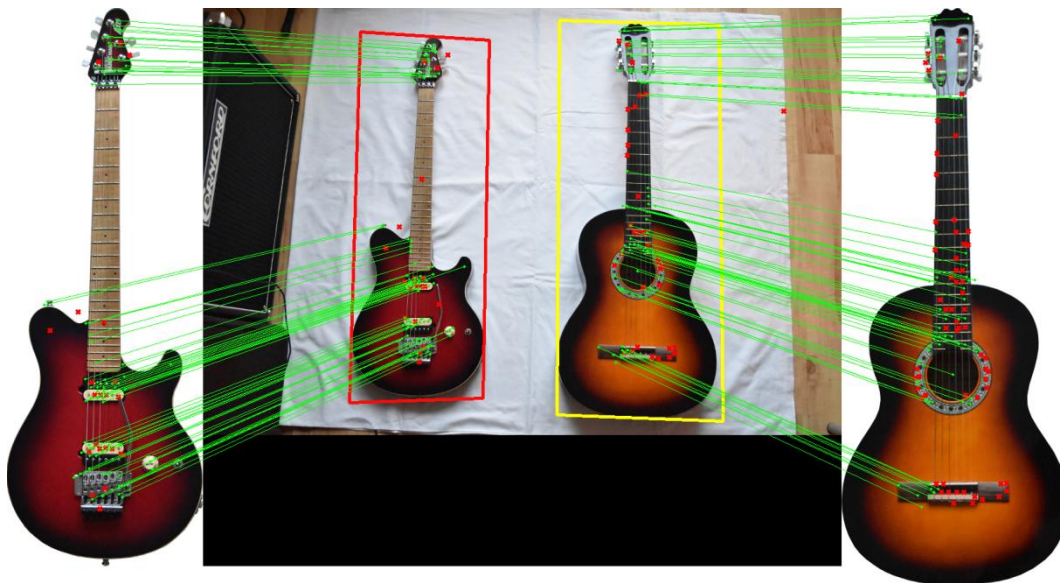
Kolejna metoda czyli **BRIEF** (ang. Binary Robust Independent Elementary Features) jest nastawiona na szybkość, dlatego nadaje się dla urządzeń o słabej mocy obliczeniowej. Rezygnuje ona z części odporności zmian skali i rotacji. Obraz zostaje poddany operacji wygładzenia i rozmycia, a następnie wybiera się zbiór punktów, które będą porównywane pewną funkcją.

Ostatnią metodą wykorzystaną przez nas w tej pracy jest **ORB** (ang. Oriented FAST and Rotated BRIEF), która jest rozszerzeniem metody BRIEF. Dodano do niej niewrażliwość na rotację. Wzorec jest poddawany serii obrotów, a następnie stosuje się metodę uczącą. Dzięki temu osiąga się większą rozróżnialność.

EKSPERYMENTY

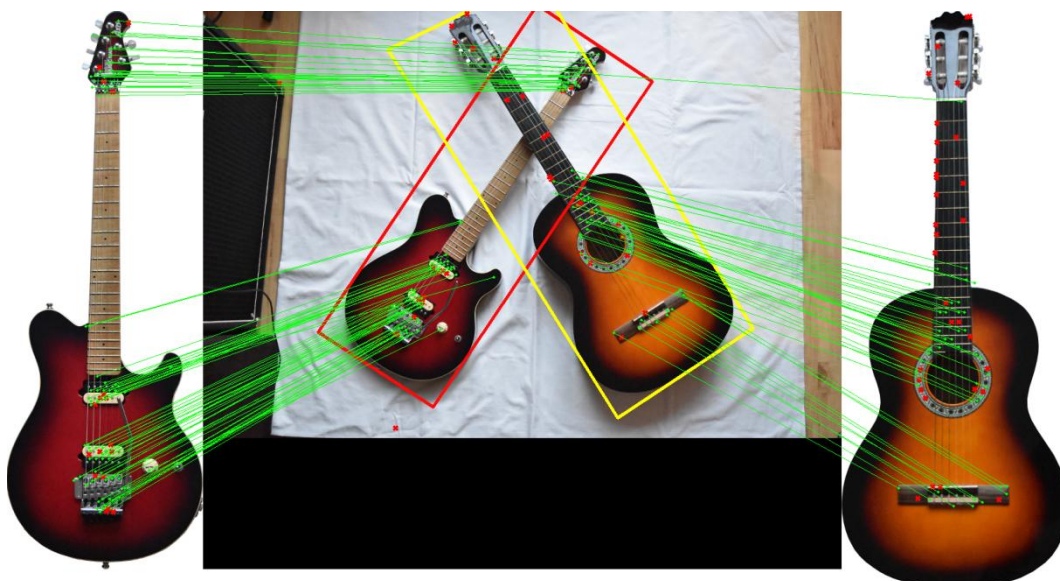
W ramach tej pracy wykonaliśmy wiele testów, aby uzyskać wiarygodne wyniki na podstawie których byłoby w stanie wyznaczyć prawidłowe wnioski dotyczące, każdego z czterech wymienionych powyżej algorytmów. W tym celu wyznaczyliśmy dwie główne kategorie obiektów, które będziemy poszukiwać, czyli gitary bez uwzględniania gryfu, oraz gitary z uwzględnieniem gryfu. Uznaliśmy bowiem, iż poszukiwanie samego korpusu będzie prostsze dla powyższych metod niż dodatkowe poszukiwanie gryfu gitary. Dla każdej z tych dwóch kategorii postanowiliśmy wyodrębnić cztery różne warunki otoczeniowe, tak aby zaobserwować jak zachowują się poszczególne algorytmy.

Należą do nich: idealne warunki, naturalne warunki ze światłem dziennym, naturalne warunki ze światłem sztucznym, oraz zdjęcia rozmyte oświetlane światłem sztucznym. Testy przeprowadziliśmy dla 100 różnych zdjęć (po 25 dla każdego warunku). Warto również podkreślić iż ustawiliśmy dosyć wysoki próg dokładności, ponieważ aby obiekt na obrazie został uznany za wykryty musi zostać znalezione 50 punktów charakterystycznych między tymi dwoma obiektami. Dzięki temu nie wykrywane są "nieistniejące gitary". Poniżej umieszczonych zostanie parę zdjęć jak dany algorytm zachowywał się przy wyszukiwaniu gitar w różnych warunkach.



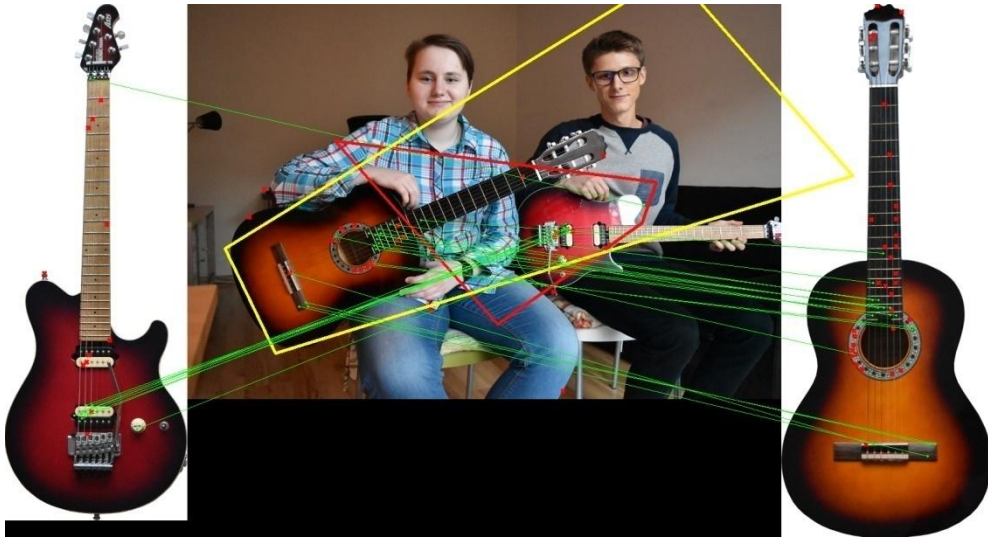
Rysunek 2 Idealne warunki, metoda SURF

Jak widzimy na powyższym rysunku, w idealnych warunkach metoda SURF, ale również wszystkie inne metody nie mają najmniejszych trudności w odnalezieniu odpowiednich gitar na obrazie. Zielonymi liniami są zaznaczone połączenia między punktami kluczowymi, które zostały odnalezione. Czerwona linia odpowiada oszacowanemu położeniu gitary elektrycznej, natomiast żółta akustycznej.



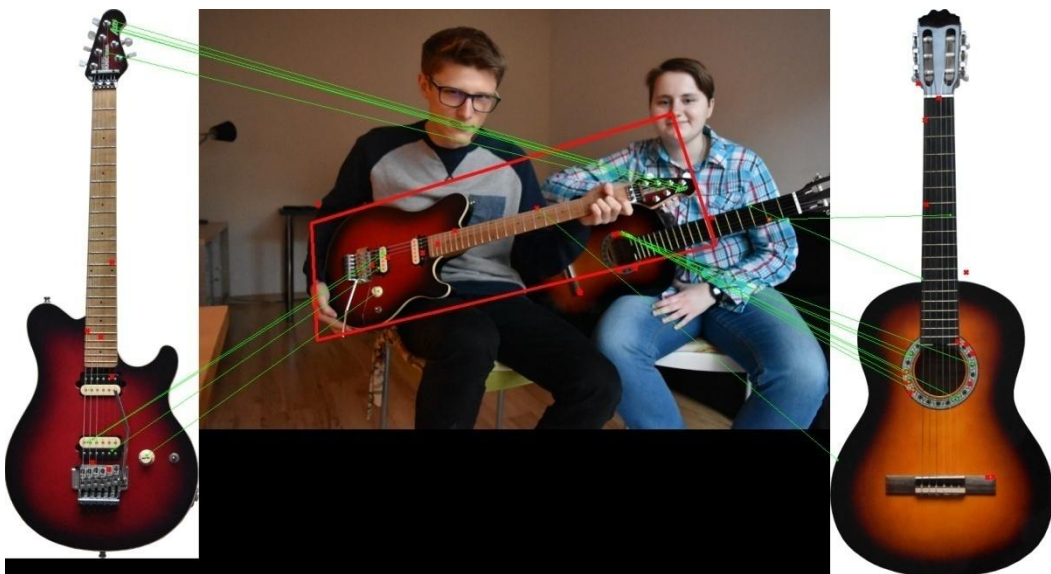
Rysunek 3 Idealne warunki + rotacja, metoda SIFT

W tym przypadku pomimo rotacji, oraz skrzyżowania obydwu gitar nie nastąpił żaden konflikt i również każda metoda określiła poprawnie położenie obydwu gitar. Należy natomiast zwrócić uwagę, że są to warunki idealne, czyli zdjęcie zostało zrobione przy naturalnym oświetleniu i przedstawione na nim gitary są pokazane od frontu (nie ma żadnych zmian perspektywy).



Rysunek 4 Normalne warunki światło dzienne, metoda BRIEF

Jak widać na powyższym rysunku, nie zawsze natomiast udaje się poprawnie określić prawidłowo położenie gitar, ponieważ każda najdrobniejsza zmiana w zmianie ich perspektywy, lub oświetleniu ma znaczący wpływ na wyszukiwanie gitar na zdjęciach.



Rysunek 5 Normalne warunki światło dzienne, metoda ORB

Widać znaczącą różnicę w ilości dopasowań pomiędzy metodą ORB, a SIFT czy SURF, dlatego też ciężko jest określić czy gitara znajduje się na danym obrazie czy też nie. Akurat w tym przypadku, gitara elektryczna została wykryta prawidłowo, natomiast dla gitary akustycznej zostało znalezionych za mało połączeń.

Powyżej przedstawione zdjęcia to tylko część testów przeprowadzonych w tej pracy. Poniżej umieszczono tabelki z wynikami z przeprowadzonych eksperymentów. W testach postanowiliśmy, że w przypadku sytuacji konfliktowych, tak jak na Rysunku 4 uznamy wynik za poprawny, ponieważ pomimo że położenie gitary nie jest określone dokładnie, to wiemy, iż na rysunku znajduje się dana gitara i jest umiejscowiona gdzieś w zaznaczonym obszarze.

WZORZEC: GITARA BEZ GRYFU

		Idealne warunki	Światło dzienne	Światło sztuczne	Rozmyte
SIFT	Elektryczna	100 %	11.76 %	0 %	0 %
	Akustyczna	100 %	35.29 %	100 %	0 %
SURF	Elektryczna	100 %	11.76 %	100 %	0 %
	Akustyczna	100 %	0 %	0 %	0 %
BRIEF	Elektryczna	33.33 %	0 %	0 %	0 %
	Akustyczna	33.33 %	0 %	0 %	0 %
ORB	Elektryczna	100 %	0 %	0 %	0 %
	Akustyczna	33,33 %	0 %	0 %	0 %

Tabela 1 Badanie efektywności danych metod przy różnych warunkach (wzorzec bez gryfu)

	Elektryczna	Klasyczna
SIFT	22.73 %	45.45 %
SURF	27.27 %	13.64 %
BRIEF	4.55 %	4.55 %
ORB	13.64 %	4.55 %

Tabela 2 Podsumowanie efektywności algorytmów dla wzorca bez gryfu

WZORZEC: GITARA Z GRYFEM

		Idealne warunki	Światło dzienne	Światło sztuczne	Rozmyte
SIFT	Elektryczna	100 %	41.18 %	100 %	100 %
	Akustyczna	100 %	23.53 %	0 %	0 %
SURF	Elektryczna	100 %	5.8 %	100 %	0 %
	Akustyczna	100 %	0 %	0 %	0 %
BRIEF	Elektryczna	66.67 %	0 %	0 %	0 %
	Akustyczna	66.67 %	0 %	0 %	0 %
ORB	Elektryczna	100 %	5.58 %	100 %	33.33 %
	Akustyczna	33.33 %	0 %	0 %	0 %

Tabela 3 Badanie efektywności danych metod przy różnych warunkach (wzorzec z gryfem)

	Elektryczna	Klasyczna
SIFT	27.27 %	4.55 %
SURF	9.09 %	9.09 %
BRIEF	54.55 %	31.82 %
ORB	22.73 %	13.64 %

Tabela 4 Podsumowanie efektywności algorytmów dla wzorca z gryfem

WNIOSKI

W czasie testów, przed badanymi metodami, postawiono wiele wyzwań. Testowano wpływ zmiany oświetlenia, występowania rozbudowanego tła czy kąt ustawienia gitar na skuteczność ich wykrywania przez poszczególne metody.

Pierwszym poczynionym przez nas spostrzeżeniem jest czas przetwarzania zdjęć przez poszczególne metody. Zgodnie z naszymi podejrzeniami, opartymi o wiadomości teoretyczne, najszybsza okazała się metoda BRIEF. Drugie miejsce zajęła metoda ORB, kolejne metoda SURF, a najwolniejsza okazała się metoda SIFT.

Na podstawie uzyskanych wyników, pomijając podział zdjęć na grupy i typ gitary, najlepsze wyniki uzyskano z wykorzystaniem metody SIFT (skuteczność wykrywania ~77%). Gorszy rezultat otrzymano w wyniku przetwarzania metodą SURF (skuteczność wykrywania ~39%). Najgorsze wyniki uzyskano metodą BRIEF (skuteczność wykrywania ~14%). Metoda ORB uzyskała wynik ~25%. Jak więc można zauważyć, kolejność metod pod względem jakości przetwarzania jest odwrotna do kolejności pod względem szybkości przetwarzania – im wolniejsza metoda, tym lepsza jakość.

Wszystkie metody otrzymywały najlepsze wyniki dla zdjęć wykonanych w idealnych warunkach (światło dzienne, jednolite tło). Zdjęcia te wykonywane były od frontu gitar. Metody SIFT i SURF nie miały z nimi najmniejszego problemu. Najgorzej wypadła metoda BRIEF, która wykryła gitary jedynie w tej grupie zdjęć ze skutecznością ~50%.

Ustawienie gitar pod różnymi kątami, dodatkowo na rozbudowanym tle, w dużym stopniu pogorszyło skuteczność wykrywania. Jak zawsze najlepiej poradziła sobie metoda SIFT. O wiele gorze wyniki zanotowały metody SURF i ORB. Metoda BRIEF nie wykryła żadnej gitary.

Zmiana oświetlenia z dziennego na sztuczne jeszcze bardziej pogorszyła wyniki.

Na zdjęciach nieostrych, gitarę wykryto jedynie w kilku przypadkach. Udało się to metodzie ORB oraz SIFT.

LITERATURA

- http://journals.bg.agh.edu.pl/AUTOMATYKA/2006-03/Auto_407-412.pdf
- http://www.ee.pw.edu.pl/~czajewsw/studenckie/magisterskie/praca_pM_WE_PW_2012_Bzdawski_Marcin_192980.pdf
- <https://repo.pw.edu.pl/docstore/download.seam%3Bjsessionid=057EE44AFE77A2A4A0EA6A15E29C45A0?fileId=WUT304267db424f4503afd9e589c849a137>
- http://docs.opencv.org/3.0-beta/doc/py_tutorials/py_feature2d/py_sift_intro/py_sift_intro.html#sift-intro
- http://docs.opencv.org/3.0-beta/doc/py_tutorials/py_feature2d/py_surf_intro/py_surf_intro.html
- http://docs.opencv.org/3.0-beta/doc/py_tutorials/py_feature2d/py_orb/py_orb.html#orb
- http://docs.opencv.org/3.0-beta/doc/py_tutorials/py_feature2d/py_brief/py_brief.html#brief