POLITECHNIKA WROCŁAWSKA WYDZIAŁ ELEKTRONIKI

KIERUNEK: Automatyka i Robotyka

SPECJALNOŚĆ: ART – Technologie informacyjne w systemach automatyki

PRACA DYPLOMOWA INŻYNIERSKA

Fotolicznik - aplikacja mobilna dla systemu iOS z przetwarzaniem obrazu

Photometer - mobile image processing application for iOS

AUTOR: Wojciech Frątczak

PROWADZĄCY PRACĘ: dr inż. Piotr Ciskowski

Spis treści

1. Wstęp 2

2. Analiza projektu 3

2.1. Warstwa aplikacji 3

2.1.1. System 4

2.1.2. Języki programowania 5

2.1.3. Architektura aplikacji 6

2.2. Rozpoznawanie licznika 7

2.3. Rozpoznawanie stanu licznika 9

3. Opis projektu 10

3.1. Lista liczników 10

3.2. Dodawanie pomiaru 11

3.3. Wizualiacja wyników 11

3.4. Informacje o projekcie 12

4. Implementacja aplikacji 12

4.1. Implementacja części bazodanowej 12

4.1.1. Technologia 12

4.1.2. Implementacja obiektów bazy danych 13

4.1.3. Zapis i odczyt obiektów 13

4.2. Implementacja interfejsu użytkownika 14

1. Wstęp

W dzisiejszych czasach należy zwrócić uwagę na szczególnie dynamiczny rozwój technologii mobilnych. Jesteśmy świadkami swoistej rewolucji mobilnej. Tablety, smartfony oraz inteligentne zegarki towarzyszą nam na każdym kroku. Wiąże się to z wieloma czynnikami. Jednym z nich jest to, że korzystanie z urządzeń mobilnych jest szybsze i łatwiejsze niż z tradycyjnych komputerów stacjonarnych lub laptopów. Następną zaletą korzystania z takich urządzeń jest możliwość wręcz błyskawicznego połączenia z Internetem w niemal każdym miejscu, oczywiście w miarę dostępu do niego za pomocą sieci komórkowej lub WiFi.

Co więcej, postęp technologiczny pozwala wykorzystywać do budowy smartfonów coraz to lepsze podzespoły i obecne urządzenia posiadają wielokrotnie większe możliwości od pierwszych urządzeń mobilnych oraz porównywalne do komputerów stacjonarnych sprzed kilku lat. Wspomniana wyżej przeze mnie popularyzacja mobilnych technologii spowodowała stworzenie ogromnej ilości aplikacji, które pomagają nam w codziennym życiu. Publikowane są one w sklepach dostępnych dla odpowiednich systemów operacyjnych. Aplikacje służą do komunikacji z znajomymi, robienia zakupów, zarządzania czasem. Prawdopodobnie dla większości, nawet najbardziej abstrakcyjnych pomysłów powstały już mniej lub bardziej udane próby implementacji aplikacji do tego służących.

Aplikacje tego typu zaczęły również powstawać w celu wsparcia zarządzania czasem, budżetem domowym, lub innymi zasobami. Jednym z takich przypadków jest zarządzanie stanami liczników wody, prądu bądź gazu, które są obecne w każdym mieszkaniu. Oprócz standardowych funkcjonalności, takich jak wprowadzanie stanu licznika, przetrzymywanie i zarządzanie nim w bazie danych, bardzo ciekawym rozwiązaniem okazuje się wykorzystanie urządzenia jako czytnika.

1. Analiza projektu

W ramach tej pracy inżynierskiej powstanie aplikacja mobilna na system iOS zarządzająca stanami liczników posiadanych przez użytkownika. Zaimplementowane zostanie przetwarzanie obrazów umożliwiające rozpoznanie konkretnego licznika a następnie rozpoznanie jego stanu. Użytkownik będzie miał również możliwość przechowywania zapisanych wartości oraz ich wizualizacji na przestrzeni czasu.

* 1. Warstwa aplikacji

Aplikacja zostanie napisana na platformę iOS. System ten jest jednym z dwóch obecnie najpopularniejszych systemów dostępnych na urządzenia mobilne obok systemu Android. Aplikacja będzie natywna, dedykowana tylko i wyłącznie na konkretny system.

* + 1. System

System iOS jest systemem operacyjnym firmy Apple dla urządzeń mobilnych iPhone, iPod touch oraz iPad. Pierwszy system, na którym bazuje obecny, został przedstawiony przez firmę z Cupertino 6. marca 2008 roku pod nazwą iPhone OS. Następnie w miarę ewolucji i udoskonalania systemu przyjął on nazwę iOS. Bazuje on na systemie operacyjnym Mac OS X 10.5 oraz jądrze Darwin.

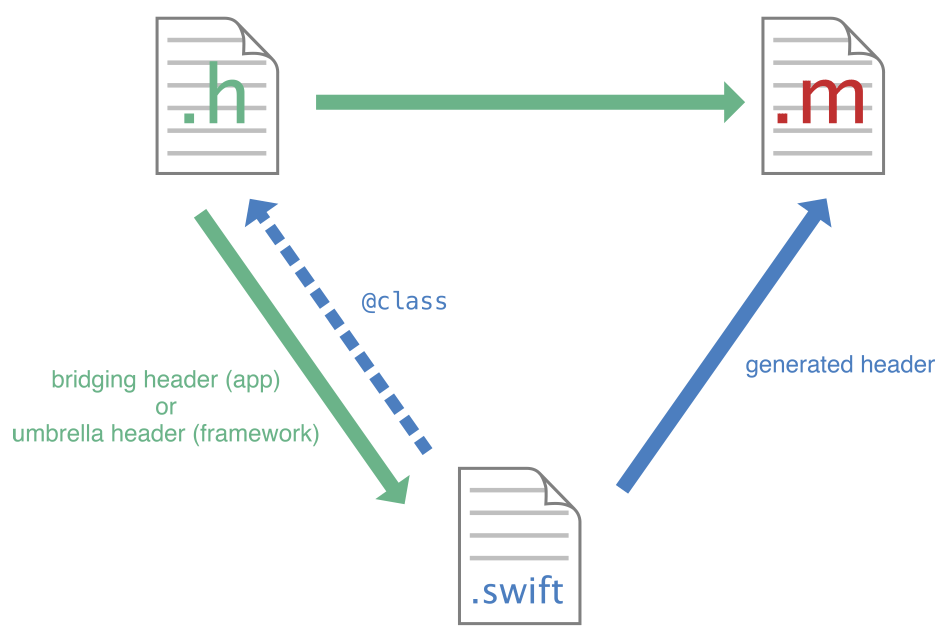
Każdy członek programu deweloperskiego może korzystać z iOS SDK w celu stworzenia swojej aplikacji. Proces publikowania tworzenia i publikowania jest ściśle kontrolowany przez firmę Apple. Do publikacji potrzebujemy posiadać odpowiednie konto deweloperskie aktywowane po uiszczeniu odpowiedniej rocznej opłaty. Deweloperzy mogą korzystać z zintegrowanego środowiska programistycznego Xcode. Dostępny jest darmowo wraz z system OS X i umożliwia tworzenie oprogramowania na systemy: OS X, watchOS, iOS oraz tvOS. Xcode posiada możliwość kompilacji kodów źródłowych języków: C, C++, Objective-C++, Objective-C oraz Swift. Program posiada również wiele ciekawych usprawnień i urządzeń ułatwiających wytwarzania oprogramowania.

System iOS składa się z czterech warstw kolejno odpowiadających za różne elementy działania, od interakcji z sprzętem po interfejs użytkownika:

* Core OS
* Core Services
* Media
* Cocoa Touch.
  + 1. Języki programowania

Aplikacja zostanie napisana w swego rodzaju hybrydzie języków programowania. Zdecydowanie największa część projektu zostanie napisana w języku Swift. Jest to nowoczesny język programowania stworzony przez firmę Apple i pokazany pierwszy raz światu podczas Worldwide Developers Conference 2 czerwca 2014 roku. Od tego momentu dość dynamicznie wypiera z rynku język Objective-C wcześniej używany do tworzenia oprogramowania na system iOS. Przez twórców jest reklamowany jako „Potężny język, który jednocześnie jest łatwy do nauki”. W mojej aplikacji wykorzystam język Swift w wersji 3.0.

Z uwagi na to, że aplikacja będzie wykorzystywać bibliotekę OpenCV napisaną w języku C++, dla potrzeb integracji języka Swift oraz C++ w jednym projekcie, część kodu zostanie również napisana w języku Objective-C. W ten sposób powstaje hybryda języków Objective-C oraz C++ nazywana Objective-C++. Pozwoliło to mi na wykorzystywanie biblioteki OpenCV wewnętrz klasy napisanej w języku Objective-C. Następnie integracja języka Objective-C z językiem Swift jest banalanie prosta przy pomocy pliku Bridging Header automatycznie stworzonego przez środowisko Xcode. W ten sposób otrzymujemy aplikację napisaną w języku Swift wykorzystującą kod napisany w języku C++.



* + 1. Architektura aplikacji

Projekt zostanie stworzony wykorzystując architekturę Model View Controller – w skrócie MVC. Jest to wzorzec architektoniczny służący organizowania struktury aplikacji. Definiuje on role obiektów a aplikacji. Wyodrębnia on trzy warstwy posiadające odpowiednio różne odpowiedzialności:

* Model – warstwa modelowa i logika aplikacji
* Widok – interfejs użytkownika
* Kontroler – komunikacja pomiędzy widokiem oraz modelem



Z korzystania z owej architektury płynie wiele benefitów i jest rekomendowana przez firmę Apple do tworzenia aplikacji.

* 1. Rozpoznawanie licznika

Aplikacja będzie miała zaimplementowaną funkcjonalność rozpoznawania licznika na podstawie dodanych do projektu wzorcowych zdjęć wszystkich liczników dostępnych w mieszkaniu użytkownika. Do przechwycenia obraz licznika wykorzystany zostanie aparat urządzenia. Jest to możliwe dzięki bibliotece AVFoundation udostępnionej dla programistów. Pozwala ona na pracę w czasie rzeczywistym z elementami audiowizualnymi smartfona, takimi jak mikrofon lub kamera.

Do przetwarzania obrazu przechwyconego z kamery posłużyłem się biblioteką OpenCV. Jest to potężna biblioteka funkcji wykorzystywanych do obróbki obrazu. OpenCV jest projektem multiplatformowym o charakterze open source.

Do rozpoznawania licznika zostanie wykorzystana metoda porównywania histogramów. Użytkownik będzie miał możliwość dodawania nieskończonej liczby liczników oraz do każdego z nich wzorcowego zdjęcia. Napisany przeze mnie program porówna przechwycony obraz za pomocą kamery do każdego zdjęcia wzorcowego. Obraz z najlepszym współczynnikiem dopasowania zostanie uznany za rozpoznany licznik. Wewnątrz aplikacji operujemy na obiektach klasy UIImage. Jest to klasa dostarczona przez framework UIKit do zarządzania danymi obrazu. Biblioteka OpenCV dostarcza nam odpowiednie metody do przetworzenia obiektu klasy UIImage na obiekt klasy cv::Mat, który z łatwością możemy przetworzyć przy pomocy OpenCV. Do porównania histogramów dwóch obrazów wykorzystałem metodę compareHist. Przyjmuje ona trzy argumenty:

* H1 – pierwszy histogram
* H2 – drugi histogram
* Metoda – metoda, którą chcemy porównać oba histogramy.

Do porównania histogramów możemy skorzystać z 4 metod:

* Metoda korelacji (correlation)



gdzie:



oraz N jest całkowitą liczbą przedziałów klasowych.

* Chi-kwadrat (Chi-square)



* Metoda przecięcia (intersection)



* Metoda odległości Bhattacharyya (Bhattacharyya distance) – OpenCV faktycznie oblicza odległość Hellingera, która jest powiązana z współczynnikiem Bhattacharyya



W swoim projekcie wykorzystam wszystkie wyżej wymienione metody. Obraz wzorcowy, który będzie miał najlepszy współczynnik w wszystkich (bądź w większości) metodach zostanie uznany za rozpoznany obraz.

* 1. Rozpoznawanie stanu licznika

Rozpoznawanie stanu licznika zostanie przeprowadzone przy pomocy biblioteki SwiftOCR. Jest to dedykowany framework na system iOS. SwiftOCR jest szybką i prostą w obsłudze biblioteką OCR (Optical Character Recognition) napisaną w języku Swift. Wykorzystuje ona również sieci neuronowe. Biblioteka jest zoptymalizowana w ten sposób, by rozpoznawać krótkie, jednoliniowe cody alfanumeryczne, więc dobrze sprawdza się do rozpoznawania stanu licznika. SwiftOCR wypada niezwykle dobrze na tle konkurencyjnego Tesseract. Autory użytej przeze mnie biblioteki dokonał testu obu bibliotek na 50 ciężkich do rozpoznania kodów alfanumerycznych. Wynik był niewzykle korzystny dla SwiftOCR.

|  |  |  |
| --- | --- | --- |
|  | SwiftOCR | Tesseract |
| Prędkość działania | 0,08 s | 0,63 s |
| Precyzja | 97,7% | 45,2% |
| Obciążenie procesora | ~30% | ~90% |
| Wykorzystanie pamięci | 45 MB | 73 MB |

Jak widzimy, SwiftOCR jest naturalnie najlepszym wyborem dedykowanym na system iOS, który wspomaga rozpoznawanie krótkich ciągów znaków.

1. Opis projektu

Aplikacja ma pełnić rolę pewnego rodzaju menadżera do zarządzania stanami liczników posiadanych przez użytkownika w mieszkaniu. Ma pozwolić na ułatwienie w regularnym spisywaniu stanu liczników oraz magazynowania ich. Wizualizacja zmagazynowanych danych pozwala na analizę i optymalizację zużycia zasobów takich jak: prąd, woda czy gaz. Aplikacja składa się z czterech zakładek, dzięki którym łatwo możemy nawigować pomiędzy widokami odpowiedzialnymi za różne funkcjonalności aplikacji.

* 1. Lista liczników

Omawiana część aplikacji odpowiedzialna jest za wyświetlenie wszystkich liczników posiadanych przez użytkownika. Lista liczników zaczerpnięta jest z lokalnej bazy danych. Oczywiście istnieje możliwość edycji każdego licznika. Możemy odpowiednio usunąć lub dodać nowy licznik do listy. Jest to możliwe za pomocą przycisków u góry ekranu. Jeśli chcemy dodać nowy licznik, konieczne jest podanie jego nazwy. Kliknięcie któregokolwiek licznika z listy przenosi nas do widoku szczegółów. Wyświetla się tam jego nazwa oraz wzorcowe zdjęcie licznika, które oczywiście możemy dodać (jeśli uprzednio nie zostało dodane) lub zmienić. Dodane tam wzorcowe zdjęcie zostanie użyte przy rozpoznawaniu licznika ze zdjęcia zrobionego kamerą urządzenia. Jest to prosta a zarazem bardzo funkcjonalna część projektu. Pozwala na przeglądanie wszystkich liczników i sprawne zarządzanie nimi.

* 1. Dodawanie pomiaru

Jest to część odpowiedzialna za najważniejszą funkcjonalność aplikacji. Pozwala ona na rozpoznanie stanu licznika i dodanie go do historii pomiarów konkretnego licznika. Zaczynamy od rozpoznanie licznika, do którego chcemy dodać pomiar. Na wstępie zostaje nam przedstawiony obraz z kamery z suwakiem odpowiadającym za przybliżenie obrazu. Po naciśnięciu przycisku „Take photo” przechwytywany jest obraz i przetworzony w celu rozpoznania konkretnego licznika. Po upływie krótkiej chwili zostaje wyświetlony komunikat z nazwą rozpoznanego licznika. Mamy możliwość potwierdzenia, lub zrobienia zdjęcia ponownie, jeśli aplikacja rozpoznała licznik błędnie. Następnie zostajemy przeniesieni do identycznego widoku, lecz posiada on już inną odpowiedzialność. Przechwytuje on obraz w celu rozpoznania już stanu licznika. Na obraz z kamery nałożony jest wąski przyciemniony pasek, który należy skierować na obszar stanu licznika. Po naciśnięciu przycisku „Take photo” aplikacja zachowuje się identycznie. Jeśli rozpoznanie zostało przeprowadzone pomyślnie, stan licznika zostaje zapisany do bazy danych przypisany do konkretnego licznika. Omawiana część aplikacji jest bardzo intuicyjna. Nie wymaga od użytkownika dużej interakcji, a jedynie dwukrotnego wykonania zdjęcia, co skutkuje na końcu dodaniem nowego rekordu do bazy danych.

* 1. Wizualiacja wyników

Trzecia część aplikacji odpowiada za wizualizację wyników. Użytkownik może wybrać jeden z listy dostępnych liczników i zobaczyć wizualizację wszystkich wyników na przestrzeni czasu. Dodatkowo, istnieje możliwość wyboru okresu czasu, którego ma dotyczyć wizualizacja pomiarów. Jest to bardzo użyteczna funkcjonalność pozwalająca na analizę zużycia odpowiednich zasobów.

* 1. Informacje o projekcie

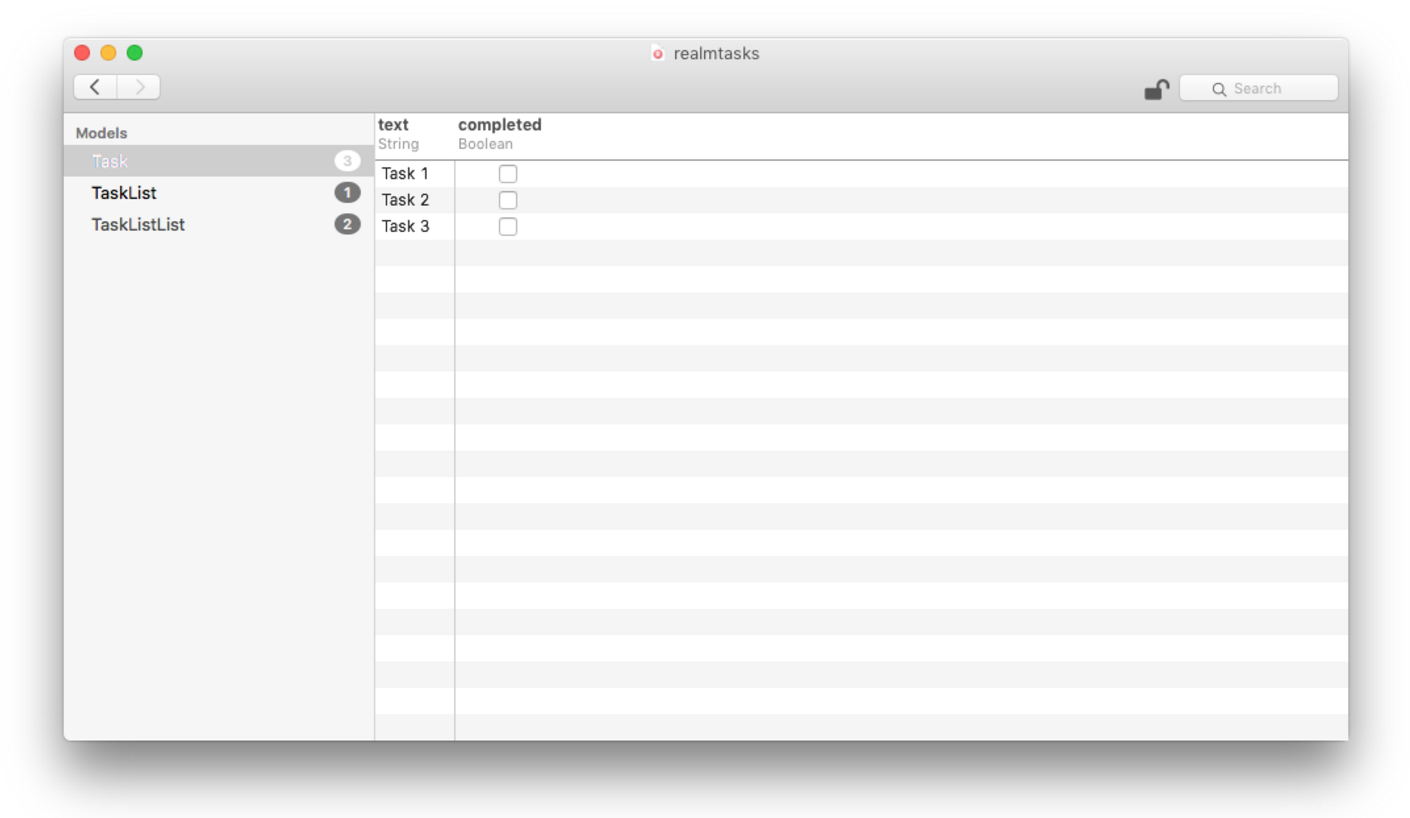
Ostatnia część aplikacji odpowiedzialna jest za przedstawienie informacji o aplikacji i jego autorze. Dodatkowo użytkownik posiada możliwość udostępnienia aplikacji na portalu społecznościowym Facebook.

1. Implementacja projektu
   1. Implementacja części bazodanowej
      1. Technologia

W aplikacji została zaimplementowana baza danych Realm. Jest to alternatywa dla powszechnie używanych baz danych Core Data, czy SQLite. Realm jest projektem typu open source. Wokół projektu powstała bardzo duża społeczność biorąca udział w jego rozwijaniu. Obecnie baza danych Realm posiada kilka milionów użytkowników. Można ją uruchomić bezpośrednio na urządzeniu, dla którego tworzymy aplikację. Realm powiada wsparcie dla bardzo wielu platform oraz języków, w tym również dla systemu iOS w językach Swift oraz Objective-C. Do głównych zalet wykorzystanej technologii należą:

* Wsparcie działania aplikacji bez dostępu do Internetu
* Szybkie tworzenie zapytań do bazy danych
* Dostęp do danych z wielu wątków
* Wsparcie dla wielu platform
* Możliwość szyfrowania danych
* Wsparcie dla programowania reaktywnego

Realm dostarcza aplikację desktopową Realm Browser, dzięki któremu możemy przeglądać całą bazę danych dostępną na urządzeniu mobilnym za pomocą interfejsu aplikacji desktopowej.

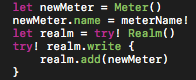


* + 1. Implementacja obiektów bazy danych

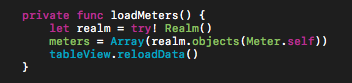
W aplikacji stworzyłem dwa rodzaje obiektów zapisywalne do bazy danych. Są to klasy Meter oraz MeterValue,   
Reprezentują one odpowiednio liczniki dodane do aplikacji oraz odczytane wartości liczników. Wewnątrz operujemy właśnie na obiektach ww. klas. W celu kompatybilności klas modelowych z bazą danych Realm muszą one dziedziczyć po klasie o nazwie Object. Jest to klasa dostarczająca możliwość zapisu oraz wczytywania obiektów z bazy danych Realm obiektów przy minimalnej ilości kodu. Każde dodane pole wewnątrz klasy dziedziczącej po klasie Object zostanie odnotowana w bazie danych bez dodatkowej implementacji.

* + 1. Zapis i odczyt obiektów

W celu zapisu (bądź wczytania) obiektów do bazy danych należy stworzyć instancję klasy Realm. Sam zapis obiektów jest bardzo prosty i sprowadza się do napisania tylko kilku linii kodu. Poniższy kod przedstawia inicjalizację nowego obiektu klasy Meter oraz jego zapis do bazy danych.



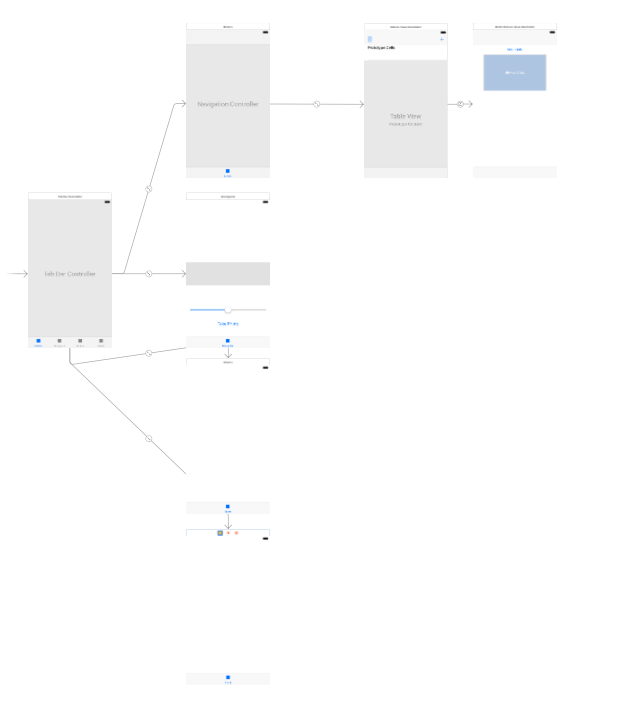
W celu odczytania obiektów z bazy danych wykonujemy równie prostą operację. Poniższy kod przedstawia prosty odczyt wszystkich obiektów klasy Meter z bazy danych.



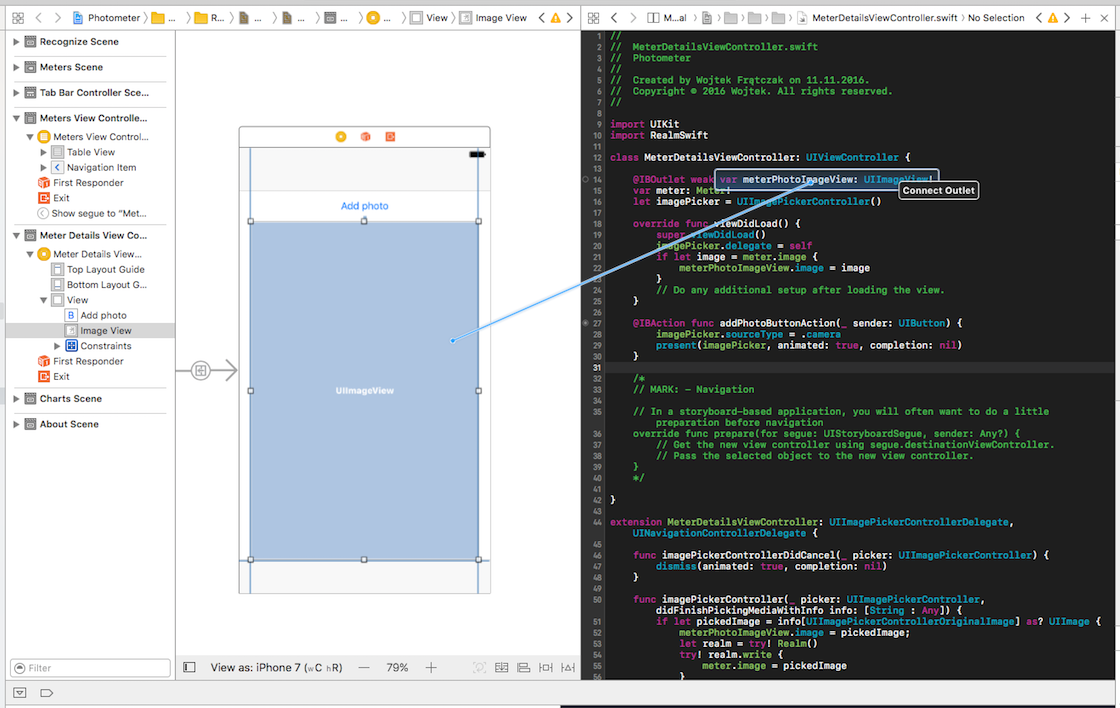
W rezultacie otrzymujemy tablicę obiektów w wyniku wywołania zaledwie jednej linijki kodu.

* 1. Implementacja interfejsu użytkownika

Do implementacji interfejsu użytkownika zostało wykorzystane narzędzie Interface Builder dostępne w ramach środowiska programistycznego Xcode. Jest to edytor, który umożliwia łatwe projektowanie interfejsu użytkownika bez wymogu pisania kodu. Technologia Cocoa Touch jest zbudowana na podstawie wzorca projektowego Model View Controller, zatem bardzo łatwe jest projektowanie warstwy interfejsu użytkownika nie ingerując w warstwy kontrolerów i modeli. Po zdefiniowaniu przynależności zaprojektowanego widoku z odpowiadającą klasą, system iOS dynamicznie tworzy połączenie między interfejsem użytkownika, a napisanym kodem. Przy pomocy plików o rozszerzeniu storyboard możemy zaprojektować poszczególne widoki oraz połączenia między nimi. W ten sposób otrzymujemy podgląd przepływu aplikacji między poszczególnymi widokami. Na poniższym obrazku został przedstawiony podgląd pliku Main.storyboard ukazujący wszystkie widoki zaimplementowane w aplikacji razem z połączeniami między nimi.

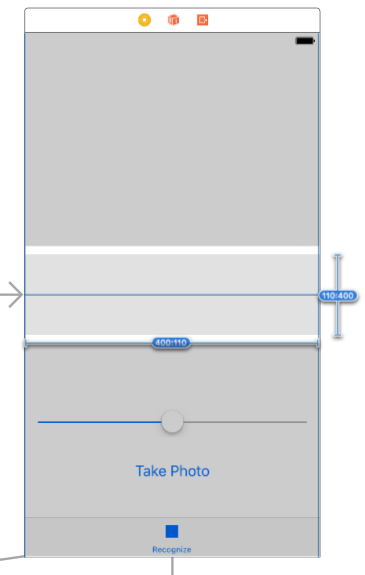


Przy budowaniu konkretnego widoku wykorzystujemy narzędzie Assistant, które pozwala nam na sprawne połączenie elementów interfejsu takich jak UILabel, UIButton i inne z kodem napisanym dla tego widoku. Poniżej przestawiam przykładowy sposób połączenia elementu z odpowiadającym polem obiektu klasy UIViewController. Analogiczną czynność wykonujemy dla połączenia akcji elementów odpowiadających za interakcję z użytkownikiem z kodem wykonywanym podczas danej interakcji, np. dla obiektów klasy UIButton.



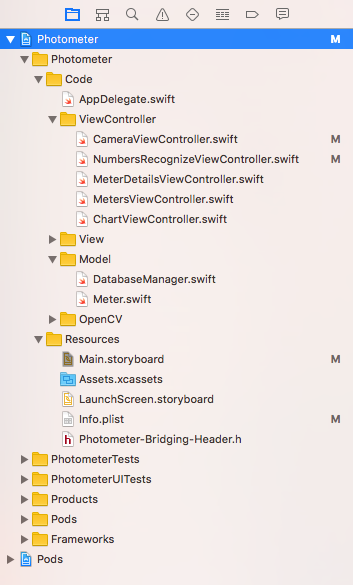
Z lewej strony obrazu widzimy strukturę zbudowanych widoków z podświetloną nazwą widoku, którym się obecnie zajmujemy oraz sam zbudowany widok. Na prawo od niego mamy dostęp do kodu napisanego lub częściowo wygenerowanego przez środowisko Xcode. Przy pomocy prawego przycisku myszy możemy sprawnie i szybko połączyć elementy interfejsu z kodem. Na obrazku zaprezentowane jest to przy pomocy niebieskiej linii.

Poszczególne elementy interfejsu umieszczamy wewnątrz wybranego widoku. Przy pomocy narzędzia Constraints możemy zdefiniować rozmiary elementu lub poszczególne odległości od innych elementów i krawędzi widoków. Na poniższym obrazku widzimy element z określonym rozmiarem i wyśrodkowaniem względem głównego widoku.



* 1. Implementacja aplikacji
     1. Struktura aplikacji

Wszystkie pliki projektu zostały podzielone na odpowiednie grupy. W środowisku Xcode pracujemy w przestrzeni roboczej, która dzieli się na dwa projekty. Pierwszy to projekt aplikacji stworzony przez nas, drugi to projekt nazwany Pods zawierający biblioteki zewnętrzne zintegrowane z naszym projektem przez narzędzie CocoaPod. Dwie główne grupy projektu naszej aplikacji to Code oraz Resources. Grupa Code zawiera wszystkie pliki kodu źródłowego podzielone zgodnie z zasadami architektury Model View Controller. Grupa Resources zawiera wszystkie inne pliki, takie jak zasoby graficzne, pliki interfejsu użytkownika, czy wspomniany wcześniej plik bridging header.



* + 1. Warstwa kontrolerów widoków

Do wcześniej zaprojektowanych widoków przy pomocy narzędzia Interface Builder stworzyłem odpowiadające im klasy kontrolerów widoków dziedziczące po klasie UIViewController dostarczonej przez firmę Apple w bibliotece UIKit. Głównym kontrolerem zarządzającym nawigacją aplikacji jest obiekt klasy UITabBarController. Klasa ta implementuje specjalny kontroler widoku, który posiada referencję do wszystkich kontrolerów widoków wchodzących w skład aplikacji. Interfejs omawianego kontrolera posiada ikonki odpowiadające każdemu kontrolerowi. Pojawia się on na dole interfejsu aplikacji. Naciśnięcie odpowiedniej ikonki kontroler przenosi nas do odpowiedniego widoku.

../../Zrzut%20ekranu%202016-11-27%20o%2021.07.41.png

Bibliografia:

https://developer.apple.com/library/content/documentation/Swift/Conceptual/Swift\_Programming\_Language/

https://developer.apple.com/library/content/documentation/Swift/Conceptual/BuildingCocoaApps/index.html#//apple\_ref/doc/uid/TP40014216-CH2-ID0

http://docs.opencv.org/2.4/modules/imgproc/doc/histograms.html

https://github.com/garnele007/SwiftOCR