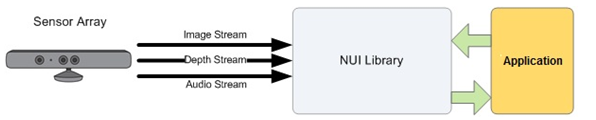
# Kinect

Urządzenie Kinect jest platformą wykrywania ruchu. Na stronie internetowej producenta można przeczytać, że Kinect for Windows "jest zestawem technologii, które umożliwiają ludziom naturalną komunikację z komputerami; czujnik oraz darmowy zestaw bibliotek i narzędzi programistycznych (SDK) Kinect for Windows dostarczają deweloperom fundament wymagany do tworzenia i rozwijania interaktywnych aplikacji, które reagują na ludzki naturalny ruch, gesty i polecenia głosowe" (tłumaczenie własne). Można zatem stwierdzić, że zestaw Kinect składa się z dwóch zasadniczych elementów - platformy sprzętowej oraz platformy programistycznej. Dane, otrzymane w wyniku przetwarzania otoczenia przez czujniki platformy Kinect, przekazywane są do aplikacji. Dzięki wykorzystaniu naturalnego interfejsu użytkownika (NUI - ang. *Natural User Interface*), sygnały wejściowe urządzenia Kinect oraz komendy sterujące programem są przyjazne i intuicyjne. Te zasady współdziałania platformy programistycznej i sprzętu przedstawia poniższy schemat.



Schemat 1.1 Interakcja sprzętu i oprogramowania z aplikacją (Microsoft Corporation, 2014)

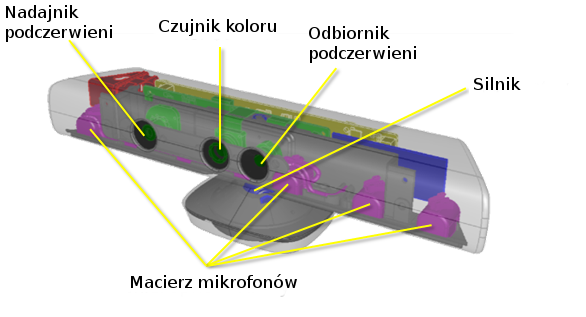
Ponadto, przydatnym narzędziem w pracy nad programami wykorzystującymi urządzenie Kinect for Windows jest Windows Developer Toolkit. Wszystkie te komponenty omówione są w podrozdziałach niniejszego rozdziału.

## Sprzęt

Urządzenie Kinect składa się z następujących podzespołów:

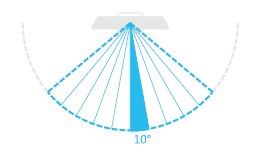
* macierz mikrofonów kierunkowych,
* promiennik podczerwieni,
* odbiornik podczerwieni,
* kamera RGB,
* akcelerometr,
* silnik

Ich rozmieszczenie obrazuje poniższy schemat:



Schemat 1.2 Podzespoły urządzenia Kinect for Windows (Microsoft Corporation, 2014)

Macierz mikrofonów składa się z czterech mikrofonów, posiada 24-bitowy przetwornik analogowo-cyfrowy. Układ wykrywa dźwięki z obszaru w zakresie pięćdziesięciu stopni na prawo i pięćdziesięciu stopni na lewo od środka czujnika. Obszar ten można podzielić na dziesięć dziesięciostopniowych fragmentów, co pozwala określić które dźwięki pochodzą z precyzyjnie określonego kierunku. Jest to rozwiązanie szczególnie przydatne do wybierania przetwarzanych dźwięków na podstawie położenia ich źródła.

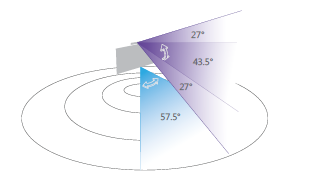


Schemat 1.3 Zakres wykrywania dźwięków (Microsoft Corporation, 2013)

Ponadto, dopuszczalne jest wprowadzenie progu natężenia dźwięku - Kinect umożliwia odfiltrowanie dźwięków poniżej 20 decybeli. Co więcej, dźwięki pochodzące zza czujnika są tłumione o 6dB. Zapewniona jest obsługa dźwięku w formacie PCM z częstotliwością próbkowania 16 kHz, z rozdzielczością 24 bitów na próbkę reprezentującym dźwięk monofoniczny (1 kanał). Z poziomu programowego dostępne są funkcje śledzenia głosu jednego ze śledzonych graczy oraz śledzenia najgłośniejszego źródła dźwięku.

Promiennik i odbiornik podczerwieni razem stanowią podsystem sensora głębokości (ang. *depth camera*), który pozwala na wykrywanie odległości pomiędzy przedmiotami i osobami znajdującymi się w zasięgu ich działania a Kinect'em. Zmierzona wartość wyrażona jest w milimetrach. Układ sensora głębokości przekazuje obraz z prędkością 30 FPS w jednej z dostępnych rozdzielczości - 320x420, 640x480 oraz 80x60. Zasięg pracy czujnika stanowi przestrzeń zakreślona między 43,5 stopniami w pionie (+/- 27 stopni względem poziomu czujnika) i 57,5 stopniami w poziomie.

Kamera RGB dostarcza obraz z taką samą prędkością jak sensor głębokości (30 FPS). Zakres pracy czujnika jest taki sam jak dla sensora głębokości. Schematycznie obrazuje to poniższy rysunek.



Schemat . Zasięg kątowy pracy czujnika głębokości i RGB

W związku z fizycznymi ograniczeniami czujników obrazu oraz głębokości, efektywne korzystanie z Kinect'a wymaga przestrzegania pewnych wymagań dotyczących obszaru działania użytkownika. Dostępne są dwa tryby działania sensora głębokości - domyślny (ang. *default*) oraz bliski (ang. *near*). Tryb działania sensora wpływa na to jakie odległości użytkownika od urządzenia Kinect są dopuszczalne i prawidłowo rozpoznawalne. W trybie domyślnym Kinect ma zakres pracy w odległości od 0.8 do 4 metrów, jednak dla najlepszej efektywności otrzymywanych sygnałów zalecane jest użytkowanie Kinect'a w odległości 1.2 do 3.5 metra od czujnika. Czujnik głębokości pracuje poprawnie także dla dystansów większych niż zalecane, ale informacje o śledzeniu i rozpoznawaniu graczy mogą być już niewiarygodne. W trybie pracy bliskiej (ang. *near mode*) zalecana odległość od czujnika zapewniająca stabilną pracę wynosi od 0.8 do 3 metrów, zaś fizycznie dopuszczalny obszar zamyka się między 0.4 a 3 metry od Kinect'a.

Podczas prac nad niniejszym projektem, gdy próbowano ustawić w programie tryb *near mode*, okazało się, że model urządzenia Kinect, którym dysponowała autorka nie jest egzemplarzem dedykowanym do użytku z systemem Windows (a co za tym idzie z pakietem bibliotek Kinect SDK) ale do użytku z konsolą Xbox. W związku z tym, tryb pracy bliskiej jest dla niego niedostępny. Poniżej przytoczono komunikat, który pojawił się podczas debuggowania błędu:

**The Kinect sensor plugged into your computer is for use on the Xbox 360. You may continue using your Xbox 360 Kinect sensor on your computer for development purposes. Microsoft does not guarantee full compatibility for Kinect for Windows applications and the Xbox 360 Kinect Sensor.**

Powyższa informacja oznacza, że urządzenie Kinect podłączone do komputera jest dedykowane do użytku z konsolą Xbox 360; dopuszczalne jest używanie go w celu rozwoju oprogramowania, jednak Microsoft nie gwarantuje pełnej kompatybilności aplikacji dedykowanych dla Kinect for Windows z czujnikiem Kinect dedykowanym dla Xbox 360 (tłumaczenie własne). W związku z tym, dla pełnej sprawności programu, zaleca się korzystanie z czujnika Kinect for Windows.

Kinect wyposażony jest także w trójosiowy akcelerometr, który może pracować w trybach 2G, 4G oraz 8G (domyślnie skonfigurowany na 2G). Obecność akcelerometru na platformie Kinect pozwala na ustalenie aktualnej orientacji układu względem kierunku grawitacji. Zaimplementowane w SDK metody pozwalają ponadto ustalić orientację względem poziomu. Daje to możliwość manipulacji kątem pochylenia zestawu Kinect for Windows za pomocą silniczka, który wmontowany jest w podstawkę urządzenia. Układ pozycjonowania Kinect'a działa z dokładnością jednego stopnia, jednak jak przestrzega producent, akcelerometr jest czuły na temperaturę pracy. W związku z tym, przekroczenie normalnej temperatury pracy (5-35o Celsjusza ) może skutkować w maksymalnie trzystopniowym odchyleniu od deklarowanej dokładności pomiaru akcelerometru (1o). Na szczęście, dla konkretnego urządzenia odchylenie to ma zawsze stałą wartość.

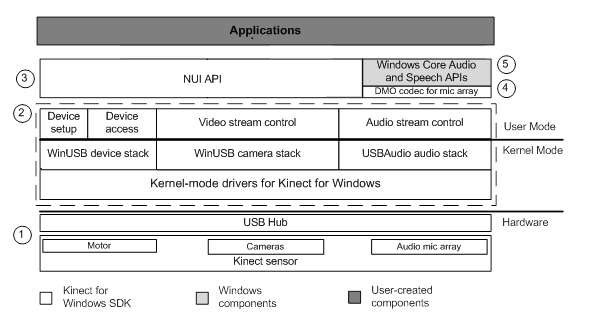
Podczas korzystania z Kinect'a w aplikacjach systemu Windows trzeba pamiętać, że ponieważ czujnik jest wyposażony jedynie w procesor sygnałowy (ang. *Digital Signal Processor*) poddający obróbce sygnał dźwiękowy z mikrofonów, przetwarzanie danych odbywa się po stronie sterownika Kinect'a na komputerze.

## Software Development Kit

Kinect for Windows Software Development Kit to zestaw pomocy do obsługi programowej czujnika na komputerach z systemem operacyjnym Windows. Zestaw ten zawiera:

* sterowniki i dokumentację techniczną,
* dokumentację i zestaw źródłowych API w języku C++, C# oraz VisualBasic,
* przykładowe projekty przedstawiające wykorzystanie dostarczonych API.

Strukturę dostarczonego SDK przedstawia niniejszy schemat:



Schemat 1.5 Architektura SDK (Microsoft Corporation, 2014)

Bloki z białym tłem oznaczają komponenty Kinect for Windows SDK, bloki z jasnoszarym tłem symbolizują komponenty systemu Windows, a bloki z ciemnoszarym tłem obrazują komponenty użytkownika (tutaj - aplikację wykorzystującą urządzenie Kinect).

Obszar oznaczony numerem jeden odnosi się do sprzętowej budowy platformy Kinect omówionej szerzej w poprzednim podrozdziale. Moduły obszaru numer dwa współpracują ze sobą w ramach sterowników urządzenia Kinect. Dzięki driverom macierz mikrofonów może być używana przez użytkownika systemu /Windows jako standardowe wejście audio. Obszar trzeci to obszar NUI (ang. *Natural User Interface*), na który składają się komponenty audio i wideo Kinect'a. Blok czwarty składa się z kodeku DMO[[1]](#footnote-1) (ang. *DirectX Media Object*) dla macierzy mikrofonów. Blok piąty to blok standardowych API Windows 7 dostępnych także dla aplikacji desktopowych Windows 8. Z powyższego schematu wynika, że czujnik Kinect wraz z dostarczonymi bibliotekami SDK tworzy całościową platformę programistyczną zintegrowaną na poziomie sprzętowym.

### Naturalny Interfejs Użytkownika (NUI)

Naturalny interfejs użytkownika jest to interfejs użytkownika, który zapewnia łatwe i szybkie nauczenie się obsługi dzięki temu, że nie jest konieczne używanie sztucznych metod i urządzeń do interakcji z aplikacją, ale wykorzystywane są naturalne metody komunikacji (np. mowa, gesty) . Programista korzystający z Kinect for Windows API buduje interfejs programu przy użyciu danych przechwytywanych przez NUI. Dane te zawierają się w strumieniu danych audio oraz w strumieniu obrazu i strumieniu głębokości. Te trzy typy sygnałów wejściowych dają możliwość stworzenia interfejsu programu przyjaznego użytkownikowi.

Na podstawie pomiarów sensora głębokości oprogramowanie obsługujące urządzenie Kinect rozpoznaje i śledzi ludzkie ciało. System jest w stanie rozpoznać nie więcej niż sześciu użytkowników jednocześnie, z czego dwóch może być aktywnie śledzonych (rozpoznawanie i śledzenie użytkowników jest szerzej omówione w kolejnych częściach pracy). Kompatybilność z Microsoft Speech API umożliwia implementację komend słownych, za pomocą których dostępne jest sterowanie aplikacją korzystającą z Kinect'a. Ponadto, dzięki dostępnym strumieniom danych zrealizowano śledzenie twarzy przy pomocy omawianego urządzenia. Wszystkie te rozwiązania dostępne są dla deweloperów w Kinect for Windows SDK.

W kolejnych sekcjach opisane są strumienie danych, z których korzystano na potrzeby pracy. Jako że nie używano przetwarzania dźwięków, nie omawiano strumienia audio. W celu głębszego poznania specyfiki tego strumienia odsyłam do dokumentacji dostępnej na stronie Microsoftu: http://msdn.microsoft.com/en-us/library/jj131026.aspx.

### Color Stream

Strumień obrazu kolorowego (RGB) umożliwia przechwycenie zwykłego obrazu z kamery Kinect'a z prędkością od 12 do 30 klatek na sekundę. Obraz dostępny jest w jednym z trzech zdefiniowanych formatów - RGB, YUV oraz Bayer. Niezależnie od wybranego formatu obraz zawsze reprezentuje te same dane, gdyż przetwarzanie do wybranego formatu odbywa się z tych samych danych źródłowych pochodzących z kamery . Podczas inicjalizacji strumienia w kodzie programu należy wybrać rozdzielczość (dostępne są 80x60, 320x240, 640x480 oraz 1280x960 pikeli) i format przechwytywanego obrazu. Tak zdefiniowanych parametrów nie można zmieniać podczas przechwytywania danych - aby je przeformułować należy wyłączyć i ponownie włączyć strumień danych obrazu kolorowego (z nowymi wartościami parametrów).

Klatka obrazu pochodzącego z kamery RGB urządzenia Kinect składa się z pikseli. Ich ilość zależy od wybranej rozdzielczości, zatem układ współrzędnych (x,y) obrazu związany jest z jego rozmiarem. Każdy piksel posiada informację o wartościach kolorów czerwonego, zielonego i niebieskiego.



Rysunek 1.1 Obraz kolorowy (źródło własne)

Strumień RGB posiada także w sobie obraz pochodzący z czujnika podczerwieni. Podczas przechwytywania strumienia można korzystać tylko z jednego obrazu na raz - albo z czujnika RGB, albo z czujnika podczerwieni.



Rysunek 1.2 Obraz podczerwony (źródło własne)

## Windows Developer Toolkit

# Bibliografia

Catuhe, D., 2012. *Programming with the Kinect for Windows Software Development Kit.* Redmond, Washington: Microsoft Press.

Microsoft Corporation, 2013. *Human Interface Guidelines v1.8.* brak miejsca:brak nazwiska

Microsoft Corporation, 2014. *Accelerometer.* [Online]   
Available at: http://msdn.microsoft.com/en-us/library/jj663790.aspx  
[Data uzyskania dostępu: 17 sierpień 2014].

Microsoft Corporation, 2014. *DirectX Media Objects.* [Online]   
Available at: http://msdn.microsoft.com/en-us/library/windows/desktop/dd375474%28v=vs.85%29.aspx  
[Data uzyskania dostępu: 17 sierpień 2014].

Microsoft Corporation, 2014. *Kinect for Windows - Meet Kinect.* [Online]   
Available at: www.microsoft.com/en-us/kinectforwindows/meetkinect/default.aspx  
[Data uzyskania dostępu: 17 sierpień 2014].

Microsoft Corporation, 2014. *Kinect for Windows Architecture.* [Online]   
Available at: http://msdn.microsoft.com/en-us/library/jj131023.aspx  
[Data uzyskania dostępu: 17 sierpień 2014].

Microsoft Corporation, 2014. *Kinect for Windows Sensor Components and Specifications.* [Online]   
Available at: http://msdn.microsoft.com/en-us/library/jj131033.aspx  
[Data uzyskania dostępu: 17 sierpień 2014].

Microsoft Corporation, 2014. *Setting Up a Kinect Sensor.* [Online]   
Available at: http://msdn.microsoft.com/en-us/library/hh855356.aspx  
[Data uzyskania dostępu: 17 sierpień 2014].

1. DMO - komponent strumieniowania danych oparty o technologię COM (tłumaczenie własne) [↑](#footnote-ref-1)