

Interakcja ludzi z „inteligentnymi” instalacjami i budynkami – główne problemy i kierunki rozwoju

Jarogniew Rykowski

1. Wstęp

Niemal od początku rozwoju „inteligentnych” instalacji interfejs człowieka i urządzeń był piętą achillesową systemu. Problem ten rozwiązywano generalnie na dwa sposoby. W pierwszym z nich maksymalnie ukrywano „inteligencję” systemu, oferując użytkownikowi interfejsy przypominające urządzenia „nieinteligentne”, na przykład klasyczny wyłącznik światła. Z drugiej strony, pozostawiano furtkę w postaci dostępu za pomocą klawiatury/ekranu albo urządzenia dedykowanego (panel sterujący np. domem inteligentnym), albo komputera przenośnego (ostatnio także tablety lub smartfony).

Każdy z tych sposobów, obok swoich zalet, ma też wady, które na dłuższą metę okazują się dość uciążliwe. Do zalet urządzeń imitujących klasyczne wyłączniki i inne elementy sterujące należy przede wszystkim wykorzystanie przyzwyczajęń pokoleń użytkowników. Ukrycie skomplikowanego układu sterującego w obudowie na przykład ściennego wyłącznika światła pozwala na sterowanie oświetleniem w sposób manualny, a jednocześnie oferuje dodatkowe możliwości, takie jak automatyczna kontrola jasności oświetlenia, symulacja obecności domowników podczas wyjazdu na wakacje, wygaszanie światła, gdy nikogo nie ma w pomieszczeniu itp. Jednakże takim układem bardzo trudno sterować – zmiana na przykład preferowanej jasności oświetlenia załączanego automatycznie wymaga albo skomplikowanych czynności manualnych (np. określonej sekwencji i czasów naciskania przycisku), albo wykorzystania centralnej aplikacji sterującej. To z kolei rodzi potrzebę zapoznania się z, często bardzo drobiazgową i nieprzyjazną, instrukcją obsługi urządzenia. Jednak kto ma czas na czytanie takiej instrukcji w celu zwykłego włączenia lampy?

Z drugiej strony obserwujemy zastępowanie klasycznego wyłącznika przez np.

ekran dotykowy sterujący oświetleniem. Niewątpliwie zakres funkcji oferowanych przez takie urządzenie jest znacznie bogatszy. Jednakże z większości tych funkcji korzysta tylko projektant systemu, który je zna, bo wcześniej przeczytał (lub nawet sam napisał) instrukcję obsługi. Typowy użytkownik nauczy się kilku podstawowych funkcji, a o reszcie nawet nie będzie wiedzieć – nie ma on czasu ani ochoty na uczenie się, jak sterować systemem. Użytkownik oczekuje wygody i maksymalnej prostoty w korzystaniu z urządzenia, a projektant i sprzedawca dążą do zapewnienia maksymalnej funkcjonalności, co niestety z prostotą użytkowania nie da się pogodzić. W rezultacie zaawansowane funkcje są wykorzystywane tylko wtedy, gdy użytkownik chce się nimi pochwalić np. przed gośćmi lub sąsiadami, w życiu codziennym są pomijane. Jeśli się jednak z nich na co dzień nie korzysta, to po co za nie płacić?

Jest zatem widoczne, że obecne interfejsy człowieka i „inteligentnej” instalacji nie spełniają wymagań ani w stosunku do oczekiwań końcowych użytkowników, ani w stosunku do wysiłku włożonego w ich projektowanie. Należy przeprowadzić gruntowną analizę potrzeb w tym zakresie, a także porównać te potrzeby z możliwościami technologii. W efekcie można zaproponować pewną strategię – nowe podejście do projektowania sposobu interakcji człowieka z instalacją i urządzeniami. Głównym celem tego artykułu jest opis takiego podejścia oraz dyskusja jego obszarów aplikacyjnych.

2. Klasyfikacja potrzeb użytkowników

Projektanci dzisiejszych systemów „inteligentnych” milcząco przyjmują dwa założenia: (1) potrzeby wszystkich użytkowników systemu są takie same i (2) należy dążyć do całościowego zaspokojenia tych potrzeb. W efekcie w systemie nie wprowadza się różnych metod interakcji

Streszczenie: Głównym celem artykułu jest przedstawienie problemu zapewnienia efektywnych metod współdziałania człowieka i urządzeń „inteligentnej” instalacji, ze szczególnym uwzględnieniem tak zwanych interfejsów naturalnych, wykorzystujących komunikację głosową, wizualną (gesty i obecność), multimedialną: wirtualną i wzbogaconą rzeczywistość itp. Głównym celem artykułu było przedyskutowanie rzeczywistych potrzeb i oczekiwań użytkowników w stosunku do możliwości technologii, a także wskazanie potencjalnych obszarów zastosowań.

Przedstawione w artykule alternatywne interfejsy z „inteligentną” instalacją wykorzystują w większym stopniu bezpośrednią komunikację z człowiekiem za pośrednictwem jego zmysłów – wzroku, słuchu oraz dotyku. Interfejsy te zakładają rezygnację z tradycyjnego tandemu klawiatura/ekran na rzecz interakcji naturalnej – rozmowy, rozpoznawania gestów i ruchów, cech charakterystycznych postaci (na przykład elementów ubioru, postawy) itp. Takie podejście przenosi ciężar nauki o użytkownika na system – to instalacja uczy się, jak najlepiej spełniać żądania i potrzeby człowieka, a nie człowiek, jak zmusić instalację do konkretnego zachowania. Dodatkowo rozbudowa instalacji staje się znacznie łatwiejsza zarówno dla projektanta, jak i przyszłego użytkownika – nowo oferowane przez system funkcje łagodnie poszerzają zakres dotychczasowych możliwości systemu, a nauka korzystania z nich nie zabiera dużo czasu.

dla różnych grup użytkowników. Rozróżnienie to następuje dopiero na poziomie urządzeń sterujących, w postaci na przykład osobnych pozycji w menu na ekranie

sterującym, ukrytych funkcji wywołanych kombinacją przycisków itp. Próbuje się w ten sposób łączyć dwa światy: codzienne użytkowanie, z założeniami jak najprostsze, oraz nadzór i naprawę błędów, z założeniami obejmujące wszystkie możliwe parametry i ich nastawy, a zatem maksymalnie skomplikowane. Tego nie da się pogodzić za pomocą jednego interfejsu. Albo urządzenie jest nieskomplikowane i proste w obsłudze, ale nie ma rozbudowanych możliwości sterowania (na przykład jest wyposażone tylko w jeden przycisk typu „włącz-wyłącz”), albo wyposażamy to urządzenie w skomplikowany, najczęściej wielopoziomowy interfejs z wykorzystaniem wielu przycisków lub nawet miniklawiatury i ekranu dotykowego. Za pomocą takiego rozbudowanego interfejsu możemy bez problemu zmienić parametry pracy urządzenia, ale tylko pod warunkiem, że umiemy się nim posłużyć. Jednakże musimy się nauczyć instrukcji obsługi nawet wtedy, gdy interakcja jest ograniczona do najprostszyc czynności (np. włączenie światła). Czyli niejako mimochodem zwyczajny użytkownik jest zmuszany do przejścia części

funkcji administratora i projektanta systemu. Dobrze, jeśli użytkownik potrafi to zrobić i uważa, że jest mu to potrzebne. Gorzej, jeśli użytkownik ten nie chce lub nie umie sobie poradzić z tak skomplikowanym systemem – naturalną reakcją będzie niechęć do korzystania nie tylko z funkcji zaawansowanych, ale także tych podstawowych. Na przykład wybór sceny świetlnej w skomplikowanym, sterowanym ekranem dotykowym wyłączniku światła odbędzie się raz i zrobi to projektant systemu. Końcowy użytkownik, nawet jeśli będzie zdawać sobie sprawę z możliwości systemu, nigdy tych ustawień nie zmieni, bo nie będzie wiedzieć jak, a nie będzie mu się chciało czytać instrukcji obsługi. Więcej, większość użytkowników przypuszczalnie w ogóle nie będzie zdawać sobie sprawy, że jest możliwa na przykład zmiana temperatury barwowej lub natężenia oświetlenia, bo projektant im tej wiedzy nie przekazał, a na samodzielne zapoznanie się z opisem możliwości systemu nie mają czasu ani chęci.

Jest zatem widoczne, że należy dogłębnie zmienić podejście do projektowania

interfejsu człowieka z „inteligentną” instalacją. Przede wszystkim należy podzielić użytkowników systemu na grupy i określić przeciętne wymagania odnośnie do sposobu interakcji dla każdej grupy. Następnie należy zaproponować efektywne metody takiej interakcji z wykorzystaniem dostępnych technologii.

Na podstawie wcześniejszej dyskusji można przyjąć, że istnieją co najmniej dwie grupy użytkowników systemu. Użytkownicy końcowi są zainteresowani maksymalną prostotą interakcji oraz minimalizacją czasu związanego z potrzebą uczenia się korzystania z systemu. Można powiedzieć, że grupa ta chce korzystać z możliwości systemu w sposób intuicyjny i maksymalnie naturalny. Natomiast administratorzy i projektanci systemu będą chcieli korzystać ze skomplikowanego interfejsu, który jednak pozwala na nieograniczoną manipulację parametrami systemu i sterowanie urządzeniami wchodzącymi w jego skład. Ta grupa z chęcią nauczy się nawet bardzo skomplikowanego sposobu obsługi, gdyż ingerencja w system jest podstawą ich utrzymania.

Wymagania obu wyżej wymienionych grup są tak skrajnie różne, że nie da się ich pogodzić w jednym urządzeniu dostępowym. Zatem należy przedyskutować możliwości technologii w zakresie urządzeń różnego typu, a następnie dobrać technologię (i sposób dostępu) do wymagań i oczekiwań danej grupy.

3. Możliwości technologii w zakresie systemów „inteligentnych”

Jak wspomniano wcześniej, w projektowaniu interfejsów z instalacją „inteligentną” dominują dwa trendy – maksymalne ukrywanie „inteligencji” pod postacią interfejsów i urządzeń sterujących przypominających rozwiązania klasyczne, oraz przeciwnie, wystawianie „na pokaz” wszystkich możliwości systemu. W pierwszym podejściu użytkownik nie musi się uczyć korzystania z systemu, ale tylko do momentu, gdy wykorzystywał jedynie jego podstawowe, czyli „nieinteligentne” funkcje. W przypadku konieczności skorzystania z którejś z zaawansowanych funkcji (na przykład określenie temperatury barwowej oświetlenia) należy skorzystać z instrukcji obsługi. Ponieważ klasyczne sterowniki nie były wyposażane w wiele przycisków, ekrany dotykowe itp., a ich rozmiary były ograniczone, o wyborze funkcji w ich nowoczesnej adaptacji decyduje często czas naciskania przycisku, sekwencja, naciśnięcie w określonym momencie w reakcji na sygnał dźwiękowy itp. Zapoznanie się z opisem tych czynności w instrukcji obsługi zdecydowanie nie jest łatwą i przyjemną lekturą. Drugi ze sposobów zakłada zastąpienie tradycyjnych urządzeń (głównie włączników) ekranem i klawiaturą z uruchomioną specjalizowaną aplikacją do zarządzania. Także i w tym przypadku użytkownik musi się mocno natrudzić, żeby nauczyć się efektywnie korzystać na przykład z rozwijanego menu lub suwaków ustalających wartości określonych parametrów. Ekran dotykowy co prawda znacznie uprościł korzystanie z wybranych funkcji, ale nie zwolnił użytkownika z konieczności nauczenia się, co, gdzie i kiedy nacisnąć. Dodatkowo niektóre implementacje interfejsu dotykowego, często nieudolnie zaadaptowane z wcześniej wykorzystywanego interfejsu komputerowego (pracującego z wykorzystaniem myszki i rozwijanych, wielopoziomowych menu), także mogą skutecznie odstraszać niezawansowanych użytkowników.

Należy także nadmienić, że o ergonomii interfejsu ekranowego decyduje najczęściej osoba, która ma wykształcenie informatyczne i często nie potrafi się wczuć w rolę niezawansowanego użytkownika – mnogość parametrów konfiguracyjnych i ich możliwych wartości nie zawsze jest pomocna dla kogoś, kto nie rozumie, czego się od niego żąda. Dla przykładu: kto potrafi określić, która „scena świetlna”, cokolwiek to znaczy, jest odpowiednia do oglądania telewizji, a jaka do czytania książki?

Niemniej istotnym problemem okazała się ilość i lokalizacja urządzeń sterujących. Z wielu względów, najczęściej ekonomicznych, urządzenia te, dostępne w niewielkiej liczbie, są lokalizowane w wybranych, często odwiedzanych miejscach (na przykład w pobliżu drzwi). Jeśli chcemy z tych urządzeń skorzystać, musimy do nich podejść, co nie zawsze jest możliwe, a na pewno nie jest wygodne. Stanie nieruchomo przy ścianie przez kilka minut wymaganych do konfiguracji parametrów urządzenia (lub zrozumienia znaczenia menu, za pomocą którego się ono z nami komunikuje) na pewno nie jest czynnością wymarzoną dla nikogo. Obserwowane ostatnio wykorzystanie w tym celu urządzeń mobilnych także nie jest idealnym rozwiązaniem, przede wszystkim ze względu na ograniczenia fizyczne telefonu/tabletu (mała i niewygodna klawiatura, kiepski ekran, niewygodne źródła zasilania itp.).

Należy też zwrócić uwagę na fakt, że jeśli urządzeń sterujących jest niewiele, to z reguły są one dość skomplikowane, bo muszą być uniwersalne – tu znowu jest widoczna potrzeba uczenia się korzystania z instalacji. Nie bez znaczenia jest także czas interakcji – na przykład, zamiast prostego naciśnięcia przycisku włączającego światło w pokoju należy przebrnąć przez kilka poziomów menu, a potem jeszcze zastanawiać się, ile lumenów zadeklarować jako jaskrawość strumienia świetlnego lampy.

3.1. Internet Rzeczy i ciche przetwarzanie danych

Lawinowy wręcz postęp miniaturyzacji urządzeń elektronicznych oraz wzrost ich „inteligencji” sprawiły, że klasyczny komputer powoli przestaje dominować jako podstawowe narzędzie pracy, rivalryki i wypoczynku. Zamiast niego obserwujemy prawdziwy wysyp niewielkich, niezauważalnych „gadżetów” elektronicznych, których głównym za-

daniem jest ułatwienie nam życia. Ponieważ gadżety takie nie rzucają się w oczy, a mimo to są bardzo użyteczne, zwykło się je określać mianem urządzeń cichego (niezauważalnego) przetwarzania danych. W ostatnich latach termin ten jest zastępowany hasłami „Internet Rzeczy” – w odniesieniu do samych urządzeń – oraz „Internet Usług” – w odniesieniu do ich funkcjonalności [1]. Urządzenia Internetu Rzeczy (i symetrycznie – usługi/funkcje Internetu Usług) mogą się łączyć w większe grupy, aby sprostać skomplikowanym zadaniom, mogą współdziałać *ad hoc* (także z ludźmi), mogą wzajemnie się wykorzystywać, mogą się przemieszczać w miarę potrzeby (czynnie – na przykład odkurzacze typu Roomba, oraz biernie – znajdując się na przykład w kieszeni nosiciela) itp.

Urządzenia Internetu Rzeczy są z reguły tak małe, że nie można ich wypożyczyć w tradycyjny interfejs znany z komputerów klasycznych – klawiaturę i ekran. W zamian należy pomyśleć o innym sposobie interakcji. Dodatkowo, skoro urządzeń ma być dużo, trudno sobie wyobrazić sytuację, w której ludzie uczą się korzystania z urządzeń – powinno być dokładnie odwrotnie, to urządzenia powinny badać i wyciągać wnioski na podstawie zachowania ludzi. Zatem należy w znacznie większym stopniu zaakcentować interakcję na zasadzie intuicji, a także metody współdziałania w trybie *ad hoc* (nieprzygotowane i nieprzewidziane z góry), poświęcając na to część „inteligencji” urządzeń. Ponieważ intuicyjna interakcja z człowiekiem musi wykorzystywać jego zmysły – wzrok, dotyk, słuch i mowę, a nawet węch – „inteligentne” urządzenia też muszą być wyposażone w protezy tych zmysłów i umieć z nich korzystać. Nie jest to już w dużej mierze problemem technicznym (co udowadnia dalsza część niniejszego tekstu), aczkolwiek pozostaje problemem organizacyjnym oraz psychicznym i społecznym.

Urządzenia Internetu Rzeczy powinny pracować zgodnie z zasadą „dobrego służącego” M. Weisera [2] – powinny być maksymalnie użyteczne przy minimalnej widoczności i stopniu „przeszkadzania” ludziom w ich codziennych czynnościach. Urządzenie powinno być zawsze gotowe do sprawowania swojej funkcji i samodzielnie (aczkolwiek w określonych ramach, narzuconych przez swojego konstruktora/właściciela) powinno decydować o swojej aktywności. Oznacza to

zmianę filozofii korzystania z urządzeń powszechnego użytku – urządzenie czeka na moment, kiedy może być użyteczne, a w efekcie samo inicjuje działanie, a nie na moment, kiedy ktoś je „wywoła do tablicy”, jawnie żądając wypełnienia pewnych funkcji.

3.2. Interfejsy naturalne

Jak wynika z wcześniejszej lektury, świadome korzystanie z urządzeń „inteligentnych” jest stopniowo zastępowane wykorzystaniem nieświadomym, intuicyjnym. Człowiek zaczyna traktować niewidzialne, ale „inteligentne” urządzenia jako coś naturalnego, jako dopełnienie sprzętów domowych i biurowych. Jednocześnie zaczyna on podświadomie domagać się takiej samej prostoty obsługi i interakcji, jak w przypadku innych inteligentnych bytów, czyli zwierząt i ludzi. O ile ten ostatni przypadek jest jeszcze ciągle w strefie marzeń, o tyle „inteligencja” urządzeń zaczyna się zbliżać do poziomu zwierząt. Mówimy tutaj o inteligencji instynktownej, a nie świadomym myśleniu i przewidywaniu (planowaniu) skutków swoich poczynań. Jednakże to właśnie instynktowne zachowanie i intuicyjny sposób komunikacji jest tym, czego zaczyna oczekiwać użytkownik. Na przykład chcemy, aby dom powitał nas tak, jak nasz pies – radosnym szczekaniem i wymachiwaniem ogonem. Oczywiście dom nie potrafi szczekać i nie ma ogona, ale mógłby w inny sposób okazać „radość” z naszego powrotu, na przykład automatycznie dostrajając oświetlenie do

naszego nastroju, przygotowując nam kawę itp. Aby osiągnąć ten cel, dom musi (1) wykryć obecność użytkownika, (2) rozpoznać go (wyróżnić spośród innych) oraz (3) znać jego potrzeby. Wszystkie te wymagania można spełnić za pomocą dzisiejszej techniki. Wykrycie obecności jest możliwe za pomocą czujek ruchu lub obecności, znanych z systemów alarmowych i z powodzeniem stosowanych od lat, a nawet prostego czujnika kontaktronowego otwarcia drzwi. Rozróżnienie użytkowników jest możliwe na przykład za pomocą analizatora obrazu z kamery, ale nie tylko – czytnik NFC [3] w drzwiach może rozróżnić wchodzących po znacznikach przypiętych do kluczy, po kodzie wpisanym w celu rozbrojenia alarmu itp. Z punktu widzenia technologii są to problemy, które już dawno zostały rozwiązane. Większy problem występuje w przypadku trzeciego z wymagań – rozpoznania preferencji i potrzeb użytkownika. „Inteligentny” dom musi być w tym zakresie samowystarczalny, to znaczy musi sam się nauczyć, czego użytkownik od niego oczekuje. Może to zrobić, obserwując jego zachowanie lub też zadając pytania. Na przykład, jeśli użytkownik od razu po przyjsciu do domu włącza telewizor i parzy kawę, czynności te można wykonać automatycznie. Co więcej, można wykryć obecność i potencjalne zamiary użytkownika, zanim on pojawi się w pokoju (na przykład wiedząc, że zostały otwarte brama i drzwi garażu, a zatem użytkownik zaraz się pojawi), co umożli-

liwia na przykład powitanie go gorącą, właśnie przygotowaną kawą.

Po pewnym czasie użytkownik przestanie zdawać sobie sprawę z faktu, że „inteligentny” dom stara się zapewnić mu maksymalny komfort i zacznie traktować taką interakcję jako coś naturalnego, podobnie jak każdy domownik traktuje radosne przywitanie ze strony psa. Podobnie, jeśli pies nas nie wita radośnie w przedpokoju, zaczynamy się zastanawiać co się stało („może jest chory?”) – potencjalny brak reakcji ze strony domu zaczniemy traktować jako symptom „choroby” i wezwiemy „lekarza”, to znaczy technika, który naprawi instalację.

Wspomniany wyżej sposób nauki potrzeb użytkownika można rozbudowywać o dodatkowe funkcje, na przykład analizę nastroju, w jakim użytkownik przyszedł do domu – na podstawie analizy twarzy z obrazu z kamery, na podstawie szybkości poruszania się po domu, na podstawie odstępstw od zwyczajowych zachowań (np. „nie usiadł w fotelu, tylko chodzi po domu”) itp. Należy wyraźnie zauważyć, że jest to stosunkowo niewielki problem technologiczny – możliwości dzisiejszej techniki pozwalają na wiele, a obiecują jeszcze więcej. W zasadzie jest to problem wystarczająco dużej wyobraźni projektanta, który musi przewidzieć w instalacji niezbędne urządzenia i napisać program, który w odpowiedni sposób przetworzy uzyskane z tych urządzeń dane.

Jak wynika z powyższej dyskusji, kluczowe staje się określenie urządzeń,

reklama

które mogą być podstawą intuicyjnej interakcji człowieka z „inteligentną” instalacją. Urządzenia te muszą komunikować się z człowiekiem za pomocą zmysłów – w równej mierze wykorzystują obraz (zmysł wzroku) i dźwięk (słuch), a także dotyk i gesty. Komunikację taką w ostatnich latach zwykło się nazywać komunikacją naturalną. Interfejsy naturalne (ang. NUI – *Natural User Interface*) zdobywają coraz większe uznanie, masowo pojawiają się nowe klasy urządzeń, które w ten sposób komunikują się z ludźmi. Poniżej zawarto opis możliwości najbardziej reprezentatywnych interfejsów naturalnych, ze wskazaniem nie tylko zalet danego sposobu komunikacji, ale także wad i przeszkód, które stoją na drodze do ich pełnego rozpowszechnienia.

Do najczęściej wymienianych interfejsów naturalnych należą analizatory (detektory) ruchu, analizatory obrazu z kamery, analizatory głosu oraz analizatory gestów.

Detektory ruchu i obecności są urządzeniami doskonale znanymi z systemów alarmowych. Urządzenia te wykorzystują promieniowanie podczerwone lub mikrofalę do wykrycia zmian rozkładu promieniowania w pomieszczeniu. Ocena wielkości zmian pozwala na określenie, czy w pomieszczeniu pojawił się ruchomy obiekt, oraz przesłanie powiadomienia o takim fakcie do centrali. Detektory ruchu są stosunkowo prostymi, a przez to tanimi urządzeniami. Jednakże ich specjalizacja w kierunku zastosowań w zakresie bezpieczeństwa sprawia, że stosunkowo ciężko jest je zaadaptować na potrzeby „inteligentnej” instalacji, z kilku powodów. Po pierwsze, podłączenie elektryczne jest realizowane w wybranym standardzie przemysłowym, które to standardy nie uwzględniają możliwości bezpośredniego dołączenia do systemu komputerowego. Zatem pobieranie i przetwarzanie danych z tych urządzeń wymaga instalacji urządzeń pośredniczących, które już nie są takie tanie, a także nie są ustandaryzowane. Po drugie, detektory nie rozróżniają poszczególnych osób, a nawet osób od większych zwierząt (bo nie miały do tego służyć – ich celem jest wykrycie intruza, a nie rozpoznanie jego tożsamości). Zatem detekcja ruchu musi być wzbogacona o system identyfikacji, na przykład za pomocą analizy obrazu z kamery. Jednakże takie wzbogacenie czyni detektor bezużytecznym, gdyż urządzenie identyfikacyjne najczęściej też ma możliwość wykrywania zmian

otoczenia i reagowania na nie. Po trzecie, detektory ruchu stosunkowo łatwo „oszukać” – eliminacja fałszywych alarmów oraz ocena zakresu i miejsca ruchu muszą być wykrywane programowo, co znowu wymaga skomplikowanego urządzenia pośredniczącego. Ze względu na powyższe wady w chwili obecnej obserwuje się stopniowe odchodzenie od prostych detektorów na rzecz detektorów znacznie bardziej skomplikowanych, na przykład analizatorów obrazu wykorzystujących kamery.

Analizatory obrazu wykorzystują obraz z kamery do oceny zmian, jakie zachodzą w pomieszczeniu, w tym analizy ruchu obiektów. Obraz może być rejestrowany w podczerwieni lub w świetle widzialnym. Obraz podczerwony służy głównie do detekcji ruchu (obecności), za jego pomocą trudno zidentyfikować jednoznacznie osobę, aczkolwiek stosunkowo prosto odróżnić np. osobę od zwierzęcia. Obraz w świetle widzialnym może być wykorzystywany do celów bardziej zaawansowanej obróbki sygnału wizyjnego, na przykład wspomnianej wcześniej analizy nastroju człowieka, emocji itp. Dane uzyskane z analizatora obrazu mogą następnie być przetwarzane w celu uzyskania konkretnych informacji, na przykład identyfikacji osoby na podstawie jej cech szczególnych, wzbudzenia alarmu (wtargnięcie nieuprawnionej osoby do strzeżonego pomieszczenia) itp. Analizatory obrazu są coraz częściej wykorzystywane do realizacji intuicyjnego interfejsu z człowiekiem, aczkolwiek ciągle wymaga to dużej pracy programistów oraz nowych technologii analizy (por. wspomniane w dalszej części tekstu analizatory gestów).

Do tej pory wskazaliśmy na popularne metody komunikacji niewerbalnej, często nieświadomej. W przeciwieństwie do nich komunikacja głosowa, czyli werbalna, polega na świadomym wypowiedaniu pewnych słów i zdań, a także na słuchaniu odpowiedzi systemu [4]. Za pomocą głosu przekazywane są: znaczenie słów (denotacja) oraz emocje, uczucia i inne wartości związane z danymi słowami (konotacja). Oprócz samego znaczeń słów bardzo ważne są też sposób i szybkość mówienia oraz intonacja, czyli sposób wypowiedania słów. Największym mankamentem komunikacji werbalnej jest konieczność rozumienia samego języka zarówno przez człowieka (użytkownika), jak i system. Ponieważ w tym procesie to system jest stroną, która jest

bardziej ograniczona, człowiek musi się nauczyć posługiwać się pewnym zbiorem języka, który dla systemu będzie zrozumiały. Zatem interfejs głosowy nie jest interfejsem w pełni intuicyjnym, gdyż trzeba mieć świadomość znaczenia wypowiedzanych słów, zdań i poleceń.

W procesie rozpoznawania mowy bardzo duże znaczenie ma komunikacja niewerbalna: wygląd fizyczny i postawa ciała, gesty, wyraz (mimika) twarzy, ruch oczu itp. Przekazywane w ten sposób informacje, często kluczowe dla zrozumienia przekazu, są pomijane przez analizatory komputerowe, które skupiają się wyłącznie na rozpoznaniu znaczenia wypowiedzanych słów. Jednakże gesty i mimika mogą przekazywać bardzo wiele informacji, od prostych komunikatów po opis głębokich emocji. Znaczenie gestów może zależeć od obszaru, kultury, społeczności, subkultury itp. Należy zwrócić uwagę na szczególną rolę ruchów rąk i dłoni. Są one zarówno elementem postawy ciała (np. ręce złożone na krzyż – pozycja zamknięta), jak i samodzielnym elementem mowy ciała. Gesty wykonywane dłońmi i rękami mogą przekazywać dużo więcej informacji, niż sama postawa ciała. Gesty dłoni można podzielić na dwie grupy. Gesty niższego rzędu to takie, które wykonywane są bez świadomości człowieka lub jako gesty automatyczne (gesty wyuczone, gesty nawykowe itp.). Gestami wyższego rzędu są gesty intencjonalne, wykonane przez osobę w celu przekazania pewnej informacji, np. wskazywanie palcem.

Do najczęściej wykorzystywanych systemów komputerowej analizy głosu [5] należą Siri, S-Voice, Google Now oraz popularna biblioteka Sphinx [6] dla języka programowania Java. Interfejsy głosowe są prawie zawsze częściami większych systemów i pełnią wyłącznie dodatkową rolę w stosunku do ich głównych funkcji. Istotną cechą wszystkich systemów automatycznego rozpoznawania mowy jest to, że są to systemy probabilistyczne, a więc nie jest możliwe opracowanie systemu o 100% skuteczności.

Jak już wspomniano, komunikacja niewerbalna odgrywa znaczącą rolę w procesie przesyłania informacji między człowiekiem i systemem. W ostatnich latach zaczęto kłaść nacisk na urządzenia, które realizują ideę komunikacji niewerbalnej, zarówno świadomej, jak i nieświadomej: narzędzia śledzenia linii wzroku (ang. *eye-tracking*) oraz analizatory mimiki i gestów.

Większość nowoczesnych systemów śledzenia linii wzroku [7] korzysta ze zmodyfikowanych kamer pracujących w zakresie podczerwieni, wykorzystujących zjawisko odbicia światła od rogówki i związaną z tym możliwość lokalizacji źrenicy oka. Systemy najnowszej generacji, w odróżnieniu od poprzednich urządzeń, nie wymagają już noszenia na głowie ciężkich hełmów z detektorami – wystarczą małe kamery ukryte w narożnikach pomieszczenia. Urządzenia na podstawie lokalizacji źrenicy oraz położenia obserwatora potrafią określić punkt przestrzeni, na który patrzy człowiek. W połączeniu z dodatkowymi urządzeniami, takimi jak analizator głosu, potrafią one zrealizować ideę intuicyjnego sterowania domem lub miejscem pracy – wystarczy popatrzeć na dany obiekt i wydać głosem polecenie, które zostanie wykorzystane do sterowania tym obiektem. Jednakże jako samodzielne urządzenia systemy śledzenia linii wzroku praktycznie nie mają zastosowania – zawsze stanowią tylko dodatek do interfejsu innego rodzaju.

W zakresie rozpoznawania mimiki i gestów najczęściej stosowanym podejściem jest wykorzystanie kamery oraz komputerowego przetwarzania obrazu [8]. Niektóre urządzenia dodatkowo wykorzystują siatkę podczerwieni do oceny głębokości sceny i odległości poszczególnych obiektów (jak np. Microsoft Kinect [9]), dwie lub więcej kamer pracujących stereoskopowo itp. Ponieważ wymagania związane z przetwarzaniem danych w takich urządzeniach są dość duże, urządzenia te mają z reguły duże wymiary lub wymagają podłączenia do komputera (np. kontroler LeapMotion [10]). Głównie ze względu na rozdzielczość kamer oraz duże wymagania odnośnie do komputerowej analizy obrazu, analizatory gestów są bardzo specjalizowane w rozpoznawaniu konkretnych gestów lub danych części ciała, którymi można takie gesty wykonywać. Na przykład Kinect, który jest samodzielnym urządzeniem o stosunkowo dużych wymiarach, pozwala na analizę gestów i postawy całego ciała z odległości kilku metrów. Kinect nie radzi sobie jednak z gestami wykonywanymi np. palcami lub z rozpoznaniem mimiki twarzy. Z kolei niewielki kontroler LeapMotion pozwala na wykrywanie gestów dłoni i palców, ale działa tylko na niewielkiej odległości (do pół metra) i nie pozwala na ocenę postawy ciała. Kontrolery gestów wbudowane w odbiorniki telewizyjne wykrywają konkretne gesty ręką, ale ich przeprogramowanie do obsługi gestykulacji nieprzewidzianej przez producenta nie jest możliwe – użytkownik musi się nauczyć, jakie gesty są rozpoznawane, co jest niezgodne z zasadą intuicyjności takiego interfejsu.

Jest widoczne, że żadna z zaprezentowanych wyżej technologii nie jest samodzielnym, pełnowartościowym interfejsem naturalnym. Jednakże technologie te bardzo dobrze wzajemnie się uzupełniają. Takie wykorzystanie wielu kanałów komunikacji naturalnej, czyli wielomodalność, jest bardzo naturalne dla człowieka, który jednocześnie mówi, gestykuluje, spogląda na otoczenie i reaguje na zmiany – na pozór chaotycznie, ale jednak w rzeczywistości w sposób skoordynowany. Zatem przyszłość należy do urządzeń wielomodalnych, które będą jednocześnie i w spójny sposób analizować mowę, gesty, być może także dotyk i linię wzroku. Niestety obecne rozwiązania są zawsze zdominowane przez jedną z technologii, podczas gdy pozostałe są traktowane jako mniej istotne dodatki. Nawet jeśli urządzenie jest wielomodalne (jak np. Kinect wyposażony w analizator mowy i gestów), to przetwarzanie danych dla różnych kanałów komunikacyjnych nie jest zsynchronizowane. Dopiero propozycja systemu, w którym wszystkie kanały komunikacji będą równoprawnie współdziałać, pozwoli utwo-

żyć pełnowartościowy interfejs naturalny między człowiekiem i „inteligentnym” systemem.

W końcowej części dyskusji na temat interfejsów naturalnych należy zwrócić uwagę na dwa istotne problemy z nimi związane: konieczność identyfikacji systemów i obiektów do sterowania w ramach „inteligentnej” instalacji oraz ochronę prywatności i nieświadome wykluczenie cyfrowe.

Problem identyfikacji urządzeń do tej pory nie był zauważany z prostego powodu – urządzeń komunikujących się za pomocą interfejsów naturalnych jest stosunkowo niewiele. Szansa, że dwa takie urządzenia będą jednocześnie wykorzystywane w tym samym czasie i miejscu, była znikoma. Sytuacja ta jednak zaczyna się zmieniać. Do tej pory w interfejsy naturalne były wyposażane tylko najdroższe urządzenia domowe, takie jak telewizory i kina domowe (sterowanie gestami i głosem) i pralki (sterowanie głosowe). Jednakże należy się spodziewać, że wraz ze wzrostem popularności interfejsów naturalnych spadną ich ceny i znacznie wzrosnie zakres zastosowań. Być może już za kilka miesięcy interfejsy te trafią do innych urządzeń, w pierwszym rzędzie do najczęściej wykorzystywanych w domu – odtwarzaczy filmów, kuchenek mikrofalowych itp. Spowoduje to sytuację, w której polecenie głosowe trafi nie do jednego, tylko do większej liczby urządzeń. Pytanie – jak wyróżnić urządzenie, do którego kieruje się polecenie? W aktualnie dostępnych urządzeniach producent milcząco zakłada, że sterowanie za pomocą naturalnego interfejsu odnosi się właśnie do tego urządzenia i nie ma ono w pobliżu „konkurentów”. Powoduje to czasami bardzo zabawne nieporozumienia (na YouTube można odszukać wiele filmów, na których np. Siri „rozmawia” z telewizorem Panasonic lub Samsung), które jednak przestaną być zabawne, gdy technologia rzeczywiście „trafi pod strzechy”. Zatem trzeba natychmiast zacząć myśleć o wzbogaceniu interfejsu naturalnego o możliwość identyfikacji urządzenia, z którym chce się przeprowadzić interakcję. Mogą to być sposoby bardzo proste – od wykrycia faktu, że stoimy przed tym urządzeniem (dotyczy urządzeń wielkogabarytowych, takich jak telewizor czy też lodówka), przez np. dotyknięcie urządzenia podczas konwersacji (co jednak ze względów praktycznych ogranicza tę konwersację do komunikacji głosowej), po zaawanso-

wane, np. śledzenie linii wzroku (i wykrycie faktu, że wpatrujemy się w telewizor). W przypadku prostych rozwiązań mogą być one uwzględnione w sprzęcie i oprogramowaniu danego urządzenia (np. wykrycie dotknięcia obudowy), ale w przypadku rozwiązań zaawansowanych trzeba wprowadzić rozwiązania systemowe, niezależne od urządzenia (trudno wyobrazić sobie niezależne systemy śledzenia linii wzroku dla każdego urządzenia i związaną z tym obecność wielu kamer w jednym pomieszczeniu). Autorowi tego tekstu nie są znane żadne rozwiązania systemowe, które idą w tym kierunku. Więcej, rozwiązania interfejsu tego samego typu, ale udostępnianego przez różnych producentów, nie są w żaden sposób ze sobą skoordynowane. Może to skutkować sytuacją, w której wydanie polecenia dla jednego urządzenia spowoduje nieoczekiwany i niepożądany efekt związany z akceptacją tego polecenia przez inne urządzenia w tym samym pomieszczeniu. Na przykład, polecenie „program pierwszy” adresowane do telewizora w łazience spowoduje rozpoczęcie cyklu prania (słowo „program” będzie rozpoznane przez pralkę jako poprawne polecenie), a polecenie „wyłącz” wyłączy wszystkie urządzenia w danym pomieszczeniu. Takie separatystyczne podejście do realizacji interfejsów przez różnych producentów praktycznie uniemożliwi sensowne wykorzystanie możliwości tej technologii, a to może spowodować brak zainteresowania końcowych użytkowników.

Drugi problem – ochrony prywatności, także jest w dzisiejszych rozwiązaniach marginalizowany przez producentów. Urządzenie, które samo ma się włączyć po wejściu użytkownika do pomieszczenia musi nieustannie śledzić wszystkich potencjalnych „klientów”. Nadzór wizyjny 24/7 we własnym domu zdecydowanie nie jest tym, co lubią użytkownicy, nawet jeśli nie do końca zdają sobie sprawę z konsekwencji tego faktu. O ile przeciętny użytkownik jest w stanie zaakceptować kamerę przy drzwiach wejściowych, o tyle kamera w salonie lub nawet w łazience jest dla większości ludzi nie do pomyślenia. Sytuację pogarsza fakt, że wyżej wzmiankowane kamery są najczęściej dostępne w sieci – kto zagwarantuje, że przechwycony obraz nie zostanie przesłany dalej, do potencjalnego włamawcy? Kto może zagwarantować, że producenci telewizorów wyposażonych w analizator gestów już tego nie robią,

choćby w celu badania rynku? Działająca kamera jest nie do odróżnienia od pasywnej, a przesyłane dane łatwo ukryć np. pod płaszczykiem aktualizacji oprogramowania. Niektórzy producenci starają się wychodzić naprzeciw tym obawom, na przykład instalując diafragmy w kamerach lub instalując je na obrotowym wysięgniku, który zawsze można odwrócić do ściany. Jednakże jak zabronić analizy głosu? Jak zdezaktywować analizator gestów, który wykorzystuje promienie podczerwone do analizy sceny i odległości od obiektów? Wystarczy, że telewizor raz dziennie prześle informację, ilu użytkowników oglądało obraz i co w tym czasie było wyświetlane na ekranie – taka informacja z punktu widzenia badania rynku byłaby bezcenna, na dodatek bardzo wiarygodna. Kto jednak da gwarancję, że to już się nie dzieje?

Ostatnim problemem związanym z masowym wykorzystaniem technologii śledzących i urządzeń Internetu Rzeczy jest nieświadome wykluczenie cyfrowe. Zjawisko to polega na tym, że nowy użytkownik, bez względu na stopień zaawansowania i obeznania z techniką, nie zdając sobie sprawy z możliwości systemu, nie będzie z nich korzystać. Na przykład osoba nieznająca kluczowych słów uaktywniających urządzenie nie będzie mogła wydawać poleceń głosowych tym urządzeniom, więcej – w ogóle nie będzie zdawała sobie sprawy z faktu, że system cały czas nasłuchuje w nadziei na wykrycie takich poleceń. Dopóki osoby tej ktoś nie poinformuje o możliwościach systemu, możliwości te pozostaną niewykorzystane (zakładając, że osoba ta postąpi jak przeciętny użytkownik, czyli nie zapozna się wcześniej z instrukcją obsługi). Problem ten może także powodować inne zjawisko – nieświadome uruchomienie urządzenia może skutkować szokiem lub co najmniej zdziwieniem, a następnie zdenerwowaniem z powodu nieumiejętności wycofania tej zmiany. Na przykład wypowiedzenie w rozmowie słów „program pierwszy” może skutkować przełączeniem kanału w telewizorze; jeśli telewizor ten nie zostanie wyposażony w klasyczny pilot, a ma wyłącznie sterowanie głosowe lub za pomocą gestów, powrót do poprzedniego kanału będzie praktycznie niemożliwy.

3.3. Interfejs namacalny i grzecznościowy

Poprzedni rozdział jest poświęcony interakcji niezaawansowanego i często

nieświadomego użytkownika z systemem. Co jednak zrobić w sytuacji, gdy system wymaga zmian, na przykład zachodzi potrzeba jego rozbudowy, zmiany parametrów działania – innymi słowy, trzeba skorzystać z usług projektanta lub administratora itp.? Interfejs naturalny jest w takim przypadku skrajnie niewygodny i zbyt wieloznaczny, aby można się było nim posłużyć. Jak zaznaczono wcześniej, interfejsy NUI nie mogą być lekarstwem na wszystkie problemy. Są one wygodne, ale zawodzą w wielu sytuacjach, gdy poziom interakcji człowieka i urządzeń staje się zbyt skomplikowany. Jest to spowodowane zbyt dużymi niejednoznacznościami w komunikacji naturalnej, zwłaszcza werbalnej. Dlatego należy pomyśleć o umożliwieniu, w pewnym zakresie i określonych sytuacjach, także interakcji za pomocą klasycznych metod typu klawiatura/myszka/ekran. Jednakże, jak już wspomniano, większość urządzeń takiej interakcji nie umożliwia. Rozwiązaniem jest tak zwany interfejs „grzecznościowy” – wykorzystanie klawiatury i ekranu innego urządzenia do przeprowadzenia interakcji w imieniu innego urządzenia lub grupy urządzeń. Na przykład zaawansowany telefon komórkowy może służyć jako pośrednik w dostępie do sterownika oświetlenia lub wentylacji. Taki sposób interakcji wymaga jednak fundamentalnych zmian w metodyce projektowania i udostępniania interfejsów, co szerzej przedyskutowano poniżej.

Interfejs „grzecznościowy” jest oparty na tradycyjnym sposobie interakcji z komputerem (klawiatura/myszka/ekran), wykorzystuje dobrze znane elementy sterujące jak rozwijane menu, przyciski ekranowe itp. Zatem jest szczególnie dobrze predysponowany do współpracy z osobami, które są przyzwyczajone do takiej formy kontaktu i uważają ją za coś naturalnego – informatykami, projektantami systemowymi, administratorami itp. Interakcja tego typu jest szczególnie efektywna w przypadku potrzeby diagnozy błędu lub testowania zachowania systemu dla różnych wartości parametrów sterujących.

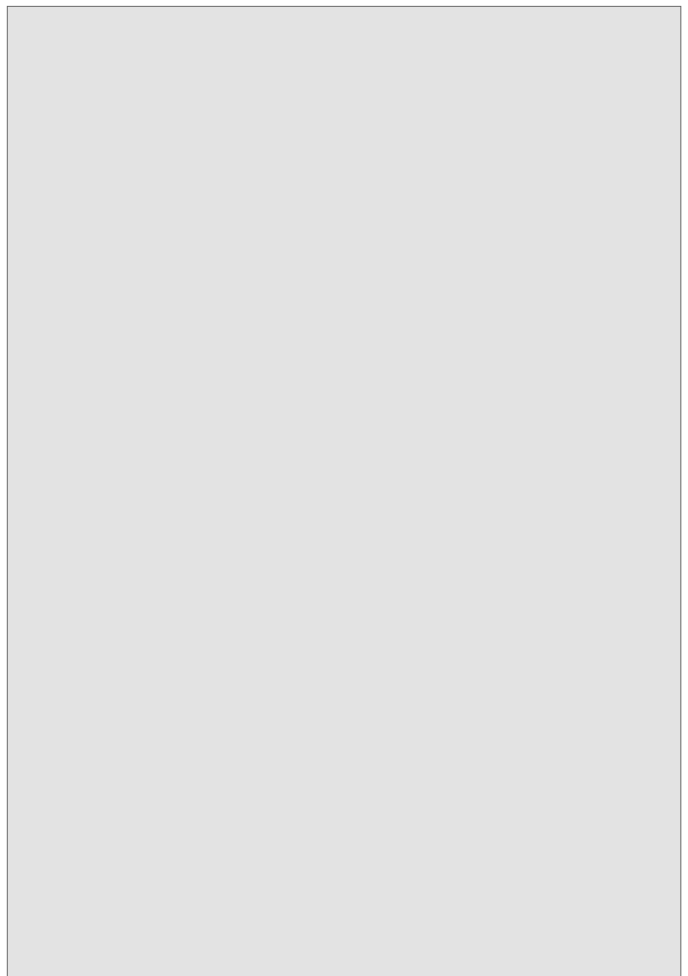
Interfejs grzecznościowy najwygodniej zrealizować za pomocą połączenia kablowego (USB) lub radiowego krótkiego zasięgu (Bluetooth lub NFC). Urządzenie wyświetlające musi być wyposażone w odpowiednie oprogramowanie. W przypadku tak zwanego „cienkiego” klienta będzie to standardowa przeglądarka internetowa, przy czym źródłem formatu i danych witryny będzie urządzenie, z którym jest nawiązywane połączenie. W trybie „grubego” klienta urządzenie to prześle do urządzenia wizualizacyjnego tylko szablon interfejsu, na podstawie którego zostanie wygenerowany sam interfejs. W pierwszym podejściu uzyskujemy dużą niezależność urządzenia pośredniczącego (wizualizującego) od danych i typu urządzenia końcowego, aczkolwiek ilość danych, które muszą być przetransmitowane, znacząco wzrasta. W przypadku połączenia z wykorzystaniem wąskiego kanału transmisyjnego (jak np. NFC) fakt ten znacząco ogranicza funkcjonalność interfejsu oraz rodzaje urządzeń, które mogą podlegać sterowaniu. W drugim przypadku ilość transmitowanych danych znacząco spada, ale odbywa się to kosztem większego wykorzystania zasobów urządzenia wizualizacyjnego (procesor i pamięć), a także wymaga zainstalowania specjalizowanej aplikacji do generowania interfejsu „w locie”. Jednocześnie można wykorzystać specyficzne cechy urządzenia pośredniczącego (np. rozdzielczość ekranu) do większej personalizacji interfejsu i dopasowania go do wymagań użytkownika oraz ograniczeń urządzenia i kanału transmisyjnego. Rosnąca „inteligencja” urządzeń powoduje, że w nowoczesnych rozwiązaniach nacisk jest przesuwany z wy-

korzystania statycznych serwerów i witryn WWW urządzeń na automatyczne generowanie interfejsu na podstawie szablonu przesłanego z urządzenia, co umożliwi większy stopień interakcji typu *ad hoc* (incydentalnej) oraz lepsze dopasowanie do specyficznych wymagań użytkownika.

Interfejs „grzecznościowy” w kontekście Internetu Rzeczy umożliwia realizację idei interfejsu „namacalnego” (ang. *palpable* [11]), czyli wykorzystywania różnych metod interakcji z urządzeniem w zależności od potrzeb. O ile standardowo użytkownik komunikuje się z urządzeniem w ramach interfejsu naturalnego, o tyle na przykład w razie awarii urządzenie musi udostępnić inny, znacznie bardziej skomplikowany interfejs, umożliwiający diagnozę i likwidację problemu. Ideę tę można rozszerzyć na zasadach kontekstu – interfejs może być dynamicznie generowany na podstawie (na przykład) praw dostępu użytkownika, jego poziomu zaawansowania, preferencji, a nawet pory dnia lub stanu innych urządzeń lub warunków zewnętrznych (na przykład automatyczne usunięcie z ekranu suwaka jaskrawości światła w środku dnia, poziomu ogrzewania w porze letniej lub gdy temperatura jest komfortowa itp.).

Interfejs „grzecznościowy” może też być z powodzeniem wykorzystywany w sytuacjach awaryjnych, czyli z definicji nieplanowanych. Na przykład stwierdzenie dymu w pokoju hotelowym może skutkować przesłaniem na telefon komórkowy planu ewakuacji oraz instrukcji postępowania w ramach pożaru, awaryjnym otwarciem drzwi itp. Podobnie urządzenia, które wykryły swoje awaryjne działanie, mogą same zainicjować proces naprawczy, przesyłając znajdującemu się

reklama



w pobliżu administratorowi prośbę o reakcję z podpowiedziami i sugestiami metod działania.

4. Wnioski końcowe

Na podstawie powyższej lektury można wysnuć pewne wnioski natury ogólnej, a na ich podstawie – sformułować postulat dla projektantów systemów „inteligentnych”. Jako pierwszy należy postawić postulat bezwzględnej rozdzielności interfejsów dla końcowego użytkownika oraz projektanta/administratora systemu. Podczas projektowania sposobu interakcji z użytkownikiem końcowym należy przede wszystkim położyć nacisk na prostotę interfejsu oraz intuicyjność jego wykorzystania, żeby uwolnić użytkownika od konieczności nauki sposobu korzystania z systemu. W tym zakresie bardzo obiecujące wydają się interfejsy naturalne, które angażują podczas interakcji najważniejsze zmysły człowieka – mowę, wzrok oraz język ciała. Mowa jest najbardziej naturalną formą komunikacji między ludźmi. Jest wykorzystywana do przekazywania informacji, wyrażania uczuć, wydawania poleceń. Komunikacja naturalna jest intuicyjnie zrozumiała niemal dla każdego, trzeba tylko poświęcić część „inteligencji” i zasobów urządzeń, żeby potrafiły się one w ten sposób porozumiewać. Rozpoznawanie poleceń głosowych, gestów, postawy ciała, mimiki itp. leży już w zasięgu możliwości technicznych dzisiejszych instalacji. Należy teraz te możliwości zacząć wykorzystywać.

W komunikacji naturalnej należy wykorzystywać jednocześnie wiele kanałów komunikacyjnych – analiza mowy powinna być połączona z analizą gestów i mowy ciała, wykrywanie ruchu i kontrola obecności – z głosowym wydawaniem poleceń (np. dla urządzeń „w pobliżu”) lub manualnym wskazywaniem urządzeń itp. Dopiero taki wielomodowy interfejs będzie użyteczny i zostanie w pełni akceptowany przez użytkowników, jako zbliżony do naturalnego sposobu porozumiewania się ludzi między sobą. Niestety, aktualne trendy w projektowaniu naturalnych interfejsów i wyposażaniu w nie urządzeń zakładają jednododalność (na przykład wyłącznie sterowanie głosem), co mocno ogranicza wykorzystanie takich interfejsów w praktyce i uniemożliwia sterowanie więcej niż jednym urządzeniem w jednym miejscu (np. pokoju).

O ile klasyczne rozwiązania – myszka, klawiatura, ekran, rozwijane menu itp. – w przypadku użytkowników końcowych się nie sprawdzają, więcej, są odbierane w większości negatywnie jako zbędny poziom skomplikowania systemu, o tyle w przypadku administratorów i projektantów systemu sytuacja jest dokładnie odwrotna. Trudno wymagać od takich osób posługiwania się interfejsami naturalnymi (np. wydawania poleceń głosowych) ze względu na niejednoznaczność i trudności w parametryzowaniu poleceń. Znacznie lepiej sprawdza się w tym przypadku interfejs klasyczny, który jest powszechnie akceptowany i używany. Trudno jednak sobie wyobrazić wbudowanie klawiatury i ekranu (lub gniazd do ich dołączenia) w każdym, najmniejszym nawet urządzeniu (na przykład regulatorem temperatury grzejnika). W zamian należy wykorzystać wyposażone w nie urządzenia personalne (smartfony, tablet) do dynamicznego generowania takiego interfejsu jako pośrednika w dostępie do urządzenia, które z kolei komunikuje się w określony sposób ze smartfonem (tabletem) na drodze radiowej. Taki interfejs, zwany interfejsem grzecznościowym, jest generowany w urządzeniu dostępowym na potrzeby chwili i na podstawie incydentalnego kontaktu ze sterowanym urządzeniem (na przykład dlatego, że takie urządzenie znalazło się w zasięgu sieci personalnej, czyli jest zlokalizowane w pobliżu komunikatora i jego użytkownika). W ten sposób za pomocą jednego, powszechnie dostępnego urządzenia dostępowego można się porozumieć z praktycznie dowolnym urządzeniem i systemem „inteligentnej” instalacji, nie wprowadzając jednocześnie zamieszania dla końcowego użytkownika, wynikającego z obecności skomplikowanego i niepotrzebnego na co dzień interfejsu.

Należy założyć, że zarówno końcowy klient, jak też projektant i administrator instalacji powinni określać sposób komunikacji na własne potrzeby – powinni mieć oni wolną rękę zarówno w wyborze sposobu interakcji, jak i w określeniu zbioru urządzeń dostępowych i pośredniczących w komunikacji. Z tego względu przy projektowaniu należy od początku uwzględnić skalowalność i rozbudowę interfejsów na życzenie oraz w zależności od sytuacji i wymagań czy też poziomu zaawansowania i wiedzy użytkownika.

Literatura

- [1] IOT-A Project – public documentation (2013), <http://www.iot-a.eu/public/public-documents>. Ostatni dostęp w październiku 2013.
- [2] Mark Weiser's home page on ubiquitous computing, <http://www.ubiq.com/weiser/>. Ostatni dostęp we wrześniu 2013.
- [3] About NFC, NFC Forum home page, <http://www.nfc-forum.org/aboutnfc/>. Ostatni dostęp w październiku 2013.
- [4] HUANG X., ACERO A., HON H.-W.: *Spoken Language Processing: A Guide to Theory, Algorithm and System Development*. 2001.
- [5] GHAI W., SINGH N.: *Literature Review on Automatic Speech Recognition*. International Journal of Computer Applications, Vol. 41, No. 8, pp. 42–50, 2012.
- [6] Strona CMUSphinx.sourceforge.net, „Basic Concepts of Speech”, <http://cmusphinx.sourceforge.net/wiki/tutorial-concepts>. Ostatni dostęp w maju 2013.
- [7] WU TUNHUA, BAI BAOGANG, ZHOU CHANGLE, LI SHAOZI AND LIN KUNHUI: *Real-time non-intrusive eye tracking for human-computer interaction*. Computer Science and Education (ICCSE), International Conference on, pp. 1092–1096, Aug. 2010.
- [8] PARKALE Y.V.: *Gesture Based Operating System Control*. ACCT'2012, pp. 318, 323.
- [9] Communicate with computers naturally, Microsoft white pages: Kinect for Windows, <http://www.microsoft.com/en-us/kinectforwindows/>. Ostatni dostęp w październiku 2013.
- [10] LeapMotion device description, <https://www.leapmotion.com/product>. Ostatni dostęp w listopadzie 2013.
- [11] Palpable Computing: A new perspective on Ambient Computing, public deliverables of PalCom, <http://www.ist-palcom.org/publications/deliverables/Deliverable-50-%5B2.7.3%5D-prototypes-final.pdf>. Ostatni dostęp w październiku 2013.

Jarogniew Rykowski – Katedra Technologii Informatycznych; Uniwersytet Ekonomiczny w Poznaniu, al. Niepodległości 10, 61-875 Poznań
e-mail: rykowski@kti.ue.poznan.pl

artykuł recenzowany