

SZEGEDI TUDOMÁNYEGYETEM
Természettudományi és Informatikai Kar
Mesterséges Intelligencia Tanszék

Szakdolgozat

Hangvezérelt Virtuális Otthon

Dávid Gergely
Programtervező informatikus

Témavezető: Dr. Gosztolya Gábor
egyetemi adjunktus

Szeged

2024.

Feladatkiírás

A program egy interaktív alaprajzon keresztül demonstrálja a beszédfelismeréssel vezérelt intelligens ház működését, lehetővé téve a felhasználók számára, hogy különböző otthoni funkciókat (például világítás, fűtés) virtuálisan irányítsanak.

Tartalmi összefoglaló

- **A téma megnevezése:**

Beszéd felismerés az okos házakban.

- **A megadott feladat megfogalmazása:**

Egy virtuális rendszerben modellezni az okos ház beszéd felismerési képességét és automatizálási rendszert.

- **A megoldási mód:**

C# programozási nyelv

- **Alkalmazott eszközök, módszerek:**

Visual Studio Community fejlesztői környezet

System.Speech beszéd felismerő könyvtár

- **Elért eredmények:**

Virtuális környezet megvalósítása, felhasználói parancsok kezelése, felhasználói visszajelzés, energiatakarékos lehetőségek beüzemelése.

- **Kulcsszavak:**

Okosház, IoT, Hangvezérlés, Mesterséges intelligencia

Tartalomjegyzék

Tartalmi összefoglaló.....	1
Bevezetés	4
Irodalmi áttekintés	5
1. Mi az okosház?.....	5
2. IoT	8
3. Hangfelismerés.....	9
Célkitűzés.....	13
Felhasznált anyagok és eszközök	14
C# Programozási Nyelv	14
Visual Studio Fejlesztőkörnyezet	15
Alkalmazott módszerek	17
Agilis szoftverfejlesztés	17
System.Speech	18
Problémák implementálásánál	19
Eredmények	21
Fonetikus írás	21
Hangvezérelt Virtuális Otthon	23
Használati útmutató:	24
További tervek	27
Összefoglalás	28
Irodalomjegyzék	29
Nyilatkozat.....	30
Mellékletek	31

Bevezetés

A szakdolgozat alapját képező irodalmi összefoglalóban áttekintem az okos házak technológiai fejlődését és a jelenlegi trendeket. Bemutatom a legfontosabb rendszereket és eszközöket, amelyek lehetővé teszik az otthonok automatizálását és vezérlését, ideértve az IoT (Internet of Things) eszközöket, a hangvezérlést és a mesterséges intelligenciát. Ezt követően egy saját okos ház rendszert terveztem és implementáltam, amely különböző eszközöket és szoftverkomponenseket integrál. Az implementáció során C# nyelvet használtam a vezérlőrendszer megírásához, valamint a Speech Synthesis API-t a hangvezérléshez. Az általam fejlesztett okos ház rendszer képes volt hatékonyan kezelni a különböző feladatokat, mint például a világítás vezérlését, a hőmérséklet monitorozását és a függönyök mozgatását. A rendszer hangvezérlése pontosnak és megbízhatónak bizonyult, a felhasználói parancsok gyorsan és pontosan végrehajtottak. Az energiafogyasztás optimalizálása érdekében a rendszer dinamikusan igazította a hőmérsékletet a megadott célértékhez, ami jelentős energia-megtakarítást eredményezett.

Az informatikai fejlesztések és az IoT (Internet of Things - dolgok internete) térhódítása új kihívások elé állítja a modern lakhatási rendszerek tervezőit és felhasználóit. Az okosházak koncepciója az emberi kényelmet, a biztonságot és az energiahatékonyságot helyezi előtérbe, lehetővé téve a lakók számára, hogy intelligens eszközök segítségével aktívan interakcióba lépjenek környezetükkel.

Az okosházak egyik kiemelkedő funkcionális eleme a hangfelismerés technológiája, amely lehetővé teszi a felhasználók számára, hogy verbális utasításokkal irányítsák otthonukat. Ez a technológia nem csupán kényelmet és hatékonyságot biztosít, hanem a hagyományos vezérlési módszerekkel szemben új lehetőségeket is teremt.

Jelen szakdolgozat célja, hogy bemutassa az okosházakban alkalmazott hangfelismerés technológia működését és alkalmazási lehetőségeit. Különös hangsúlyt fektetünk egy saját fejlesztésű C# programra, amely lehetővé teszi a lakás hang-vezérelt irányítását. A program tervezése és implementációja során különös figyelmet fordítottunk a felhasználói élményre és az energiahatékonyságra.

Irodalmi áttekintés

1. Mi az okosház?

Az okos otthon vagy okosház olyan lakóépület, amelyben a különböző elektronikai és informatikai eszközök összekapcsolódnak egy központi hálózaton keresztül azért, hogy automatizálják és optimalizálják a lakók kényelmét, biztonságát és energiafelhasználását [1]. Az okosházak számos érzékelőt, eszközt és vezérlőt használnak, amelyek lehetővé teszik a lakók számára, hogy távolról is irányítsák az otthonukat, például okostelefonról vagy táblagépről.

Ez a leírás azt is takarja magában, hogy nincs egy egyértelmű elválasztás, hogy mit nevezünk okosháznak és mit nem, általában ez csak egy hangzatos marketingfogás, hogy eladjunk pár egyszerű IoT terméket egy csomagba.

Azért, hogy jobban megismerjem, hogy mi is egy mostani vásárlónak az igénye szétnéztem a digitaltrends oldalán [2] specifikusan okos eszközöket keresve a lakásba. Jól látható, hogy a nyár közeledtével megszaporodtak az okos klíma berendezéseket hirdető szövegek száma, ahol többféleképpen próbálják a vásárlókat elérni. Van akit azzal csábítanak, hogy telefonról bárhol tudja irányítani a készüléket, így hűvös a lakás mire hazaér, másoknak a környezettudatos oldalát fogják meg, jellemzően azzal, hogy az eszköz nem egy meghatározott teljesítményen megy folyamatosan, hanem igazodik a környezethez így képes kisebb teljesítménnyel hasonló eredményt elérni és később romlik el. A drágább kategóriákat az különbözteti meg a többitől, hogy ezek már hordozhatóak. Ezek az eszközöket már nem is lehet légkondicionálónak nevezni, hanem (air purifier) légtisztítónak nevezik annak ellenére, hogy képesek hűtésre és fűtésre egy Dyson Cool TP07 544\$-ba azaz majdnem kétszázezer forintba kerül.

A biztonsági rendszerek az okosházak fontos elemei közé tartoznak, amelyek célja a lakók biztonságának és nyugalmanak megőrzése. Az okosházakban alkalmazott biztonsági rendszerek számos eszközt és technológiát használnak, amelyek közül az egyik legfontosabb a kamerarendszer [4]. A kamerák lehetővé teszik a lakók számára, hogy folyamatosan

figyelemmel kísérik otthonuk környezetét, és megfigyelik a különböző területeket és bejáratokat.

Az okosházak biztonsági kamerarendszerei általában magas felbontású kamerákból állnak, amelyek éjjel-nappal képesek kiváló minőségű felvételek készítésére. Ezek a kamerák gyakran infravörös technológiát használnak, amely lehetővé teszi a látás javítását rossz fényviszonyok között is. Emellett a legtöbb okosház biztonsági rendszere intelligens funkciókkal is rendelkezik, például mozgásérzékelőkkel és arc- vagy alakfelismerő szoftverrel, amelyek automatikusan riasztást küldenek a lakóknak, ha gyanús tevékenységet észlelnek.

Az okosházak kamerarendszerei távoli hozzáférést is biztosítanak a felhasználók számára, így a lakók bármikor és bárhol ellenőrizhetik otthonukat egy okostelefon vagy egyéb internetkapcsolattal rendelkező eszköz segítségével. Ezáltal a kamerák nemcsak a biztonságot növelik, hanem a lakók nyugalma és biztonságérzetét is fokozzák.

Az okosházak biztonsági kamerarendszerei tehát fontos szerepet töltenek be az otthoni biztonság és a felhasználók kényelme szempontjából, biztosítva, hogy a lakók mindig tisztában legyenek otthonukban zajló eseményekkel és tevékenységekkel.

Az okosházak egy másik fontos eleme az AI asszisztensek, mint például Siri és Alexa, amelyek lehetővé teszik a felhasználók számára, hogy verbális utasításokkal irányítsák és kezeljék otthonukat. Az AI asszisztenseknek számos hasznos funkciójuk van az okosházakban, például az eszközök vezérlése, az információk lekérése, az emlékeztetők és naptárbejegyzések kezelése, valamint egyéb otthoni automatizált folyamatok irányítása.

Az AI asszisztensek integrálhatók az okosház rendszerébe, és képesek kommunikálni más okos eszközökkel és rendszerekkel a házban. Például Siri és Alexa képesek vezérelni a világítást, a fűtést, a klímaberendezéseket, az okos dugaljakat és sok más eszközt az otthonban.

Ezek az AI asszisztensek gyakran rendelkeznek természetes nyelvfeldolgozási képességekkel, ami lehetővé teszi számukra, hogy értelmezzék és reagáljanak az emberek által beszélt utasításokra és kérdésekre. Ezáltal a felhasználók egyszerűen és intuitívan kommunikálhatnak az okosházukkal anélkül, hogy szükség lenne bonyolult interfészekre vagy beállításokra.

Siri és Alexa továbbá képesek integrálni más okos otthoni alkalmazásokat és szolgáltatásokat, például zene lejátszást, hírek és időjárásinformációkat, valamint online vásárlási lehetőségeket. Ezáltal az AI asszisztensek széles körű funkcionalitást és szórakoztató lehetőségeket kínálnak a felhasználók számára az okosházakban.

Az AI asszisztensek, mint Siri és Alexa, így központi szerepet töltenek be az okosházakban, lehetővé téve a felhasználók számára, hogy egyszerűen és hatékonyan irányítsák és kezeljék otthonukat a mesterséges intelligencia által nyújtott segítséggel.

Ezek az okos eszközök rendkívül előnyösek lehetnek különböző hátrányos helyzetű embereknek is például az interfészek helyetti hangfelismerő vezérlésre egy látássérült is képes anélkül, hogy órákig gyakorolná, hogy egy kijelzőn miután mit kell nyomni még a zárat is át lehet alakítani úgy , hogy telefonnal nyithatóak legyenek így nem kell a kulcsok között keresgélnie, vagy a redőnyök automatikus le és felhúzását mind meghatározott időben mind hangvezérléssel sokkal könnyebb használnia egy tolószékesnek vagy bármilyen más testi fogyatékkal élőknek. Szintén automatikus világítással segíthetünk mozgássérülteknek, hogy ne kelljen a kapcsolót használnia. Ezen felül akár időseknek akár hátrányos helyzetűeknek tud egy biztonsági háló lenni az, ha például egy smart watch küld egy értesítést szeretteinknek, ha valami rendellenességet tapasztal (például alacsony pulzus vagy egyéb abnormális rendellenesség) és ezután azonnal meg is tudják nézni a beépített kamerákon, hogy valójában mi is történik.

Ezzel a kamerás rendszerrel át is térhetünk a biztonságra. Természetesen szép és jó, hogy okos eszközökkel képesek vagyunk kinyitni az ajtókat viszont és le tudjuk húzni a függönyöket vagy még az ablakokat is kinyitni, viszont, ha egy lakás ilyen eszközökkel van felszerelve az vonzhat olyan rosszakarókat is akik hajlamosak extra erőfeszítést tenni hogy kirabolják a házat. Ennek elkerülése érdekében több védekezést is alkalmazhatunk és mivel egy megoldást könnyebb áttörni ezért érdemes a lehetőségeink kombinációjával védenünk otthonunkat. Fokozottan kell ügyelnünk arra, hogy minden rendszernél a legtöbb gyártó által biztosított védelmet használjuk mint például a többlépcsős verifikációk a jelszavakra, de ezen felül érdemes rendszeresen megváltoztatni a jelszavakat, ebben egyébként több előfizetéses szolgáltatás is tud segíteni, például a Nord VPN-nél is van erre lehetőség.

2. IoT

Az Internet of Things (IoT), vagyis a Dolgok Internete egy dinamikus és gyorsan fejlődő technológiai paradigmaváltás, amely forradalmasítja az emberek mindennapi életét és az iparágak működését [4]. Az IoT lehetővé teszi a különféle eszközök és rendszerek intelligens összekapcsolását és kommunikációját az interneten keresztül, megteremtve ezzel az okos és összekapcsolt környezeteket. A hagyományos eszközök, például háztartási készülékek, gépek és érzékelők, valamint az újabb okoseszközök, mint például okosórák és okostelefonok mind az IoT részét képezik, hozzájárulva a mindennapi tevékenységek automatizálásához és optimalizálásához.

Az IoT egyik legnagyobb előnye az adatok hatalmas mennyisége, amelyet az eszközök és érzékelők folyamatosan gyűjtenek és megosztanak az interneten keresztül. Ezek az adatok lehetővé teszik az okos rendszerek és alkalmazások számára, hogy valós időben reagáljanak a környezeti változásokra és az emberi viselkedésre, elősegítve a hatékonyabb energiafelhasználást, a jobb teljesítményt és a felhasználói élmény javítását. Például az okos otthoni rendszerek felhasználhatják az érzékelők által gyűjtött adatokat a fűtés és hűtés optimalizálására, az energiafogyasztás csökkentésére és az otthoni biztonság növelésére.

Ez a technológia az ipari szektorban is jelentős változásokat hoz, lehetővé téve a gyártók számára, hogy okosabbá tegyék a termelési folyamatokat és növeljék a

termelékenységet. Az ipari IoT eszközök és rendszerek lehetővé teszik az üzemek számára, hogy valós időben monitorozzák és optimalizálják a gépek teljesítményét, elősegítve a karbantartást és a javításokat, valamint minimalizálva az üzemzavarokat és leállásokat.

Az IoT azonban számos kihívással is szembesül, különösen a biztonság és adatvédelem terén. Az IoT eszközök és rendszerek nagyszámú kapcsolódási pontot jelentenek, amelyek potenciálisan sebezhetővé tehetik őket a kiberbűnözés és a támadások számára. Ezért fontos, hogy az IoT tervezése és implementációja során prioritást kapjon a biztonságos tervezés és a szigorú adatvédelmi intézkedések bevezetése.

Összességében az IoT ígéretes jövőt kínál a technológiai fejlődés és az innováció szempontjából, lehetővé téve az intelligens és összekapcsolt világ létrehozását, ahol az eszközök és rendszerek egyszerűbben és hatékonyabban kommunikálhatnak egymással és az emberekkel, hozzájárulva ezzel a fenntartható és kényelmes jövőhöz.

3. Hangfelismerés

Első lépés a hangfelismeréshez, hogy számítógéppel értelmezhető formára hozzuk a beszédet. Ez gyakorlatilag egy analóg jel digitalizálását jelenti, ami azt jelenti, hogy valamilyen mintavételezéssel adatokat kell kinyernünk a hangból [5]. Jellemzően ezt a folytonos időben változó jelet diszkrét időközönként mérjük.

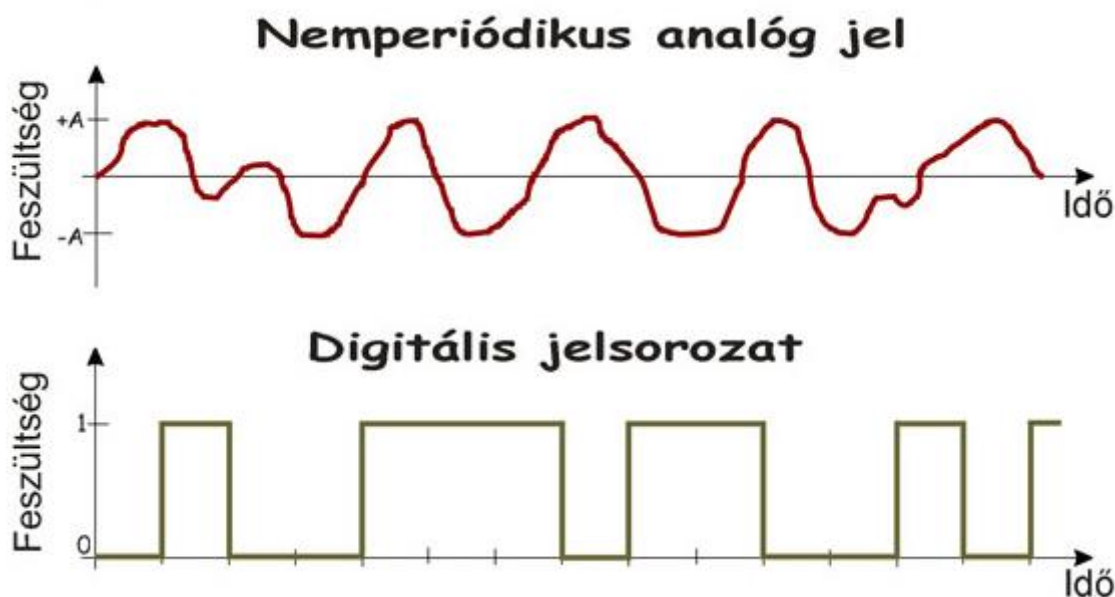
A mintavételi frekvencia:

$$f_s = \frac{1}{T}$$

Ahol T a mintavételi időköz ez adja meg, hogy másodpercenként hány mintát veszünk.

Miután tudjuk, hogy milyen gyakran veszünk mintát a jel amplitúdóját is diszkrétizáljuk. Mivel nem tudunk tökéletesen mintát venni ezért meg kell elégednünk egy „elég jó” megoldással, a számítógép nem képes végtelen mennyiséget kezelni ezért a felvett hangok jelszintjét egy véges halmazba korlátozzuk. 8 bites kvantálás esetén -128 és 127 közötti értékeket míg 16 bites kvantálás esetén -32768 és 32767 között lehetnek, míg amik ezen túllógnak egyszerűen levágjuk. Így az amplitúdó kvantálás után a jel mindkét tengelyen

diszkrét lesz és a jel digitális lesz. Természetesen ez a megoldás sose fogja tudni tökéletesen reprezentálni a felvett hangot, különösen, ha a jel nagyobb dinamikatartományú, mint amit a kvantálás képes kezelni, ezek a kilengő adatok miatt torzulhatnak az adataink, ami hibás beszédfelismerést eredményezhet. Az emberi fülhöz körülbelül 20 bit mélységű mintavételezésre van szükség, hogy minden hangot lefedjünk, amit az ember képes hallani, de beszédfelismerésre elég csak 16 mert kevesebb hangot tudunk képezni, mint amennyit hallunk (Mintavételezési tétel). Erre egy példa az 1.ábrán.



1.ábra: Analóg jel digitálissá váltása

<https://www.szabobarnabas.com/1-digital/analog>

Miután a hangjelet digitalizáltuk a következő fontos lépés az előfeldolgozás, amely során a nyers hangjelet tisztítjuk és előkészítjük a felismerési folyamatra. Ez a szakasz kritikus, mert a nyers hangjelek gyakran tartalmaznak zajt és egyéb zavaró tényezőket, amelyek akadályozhatják a pontos felismerést. Az előfeldolgozás több kulcsfontosságú lépést tartalmaz.

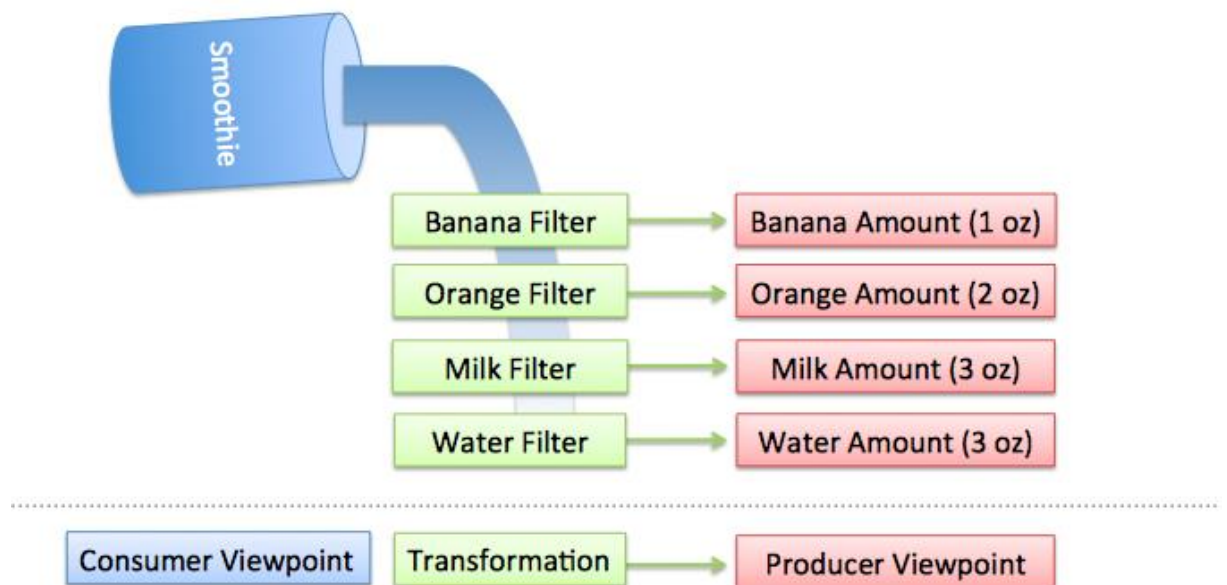
Az első lépés a zajszűrés, amely során a háttérzajt eltávolítjuk a hangjelből. A háttérzaj lehet például egy ventilátor zúgása, közlekedési zajok vagy egyéb nem kívánt hangok, amelyek torzítják a beszédet. A zajszűrés során különböző eljárásokat alkalmazhatunk, mint például az adaptív szűrést, amely figyelembe veszi a zaj jellemzőit és dinamikusan alkalmazkodik hozzá.

Ez után a hangjelet szegmentáljuk, kisebb részekre bontjuk például szótagokra vagy fonémákra. Ez lehetővé teszi, hogy a beszédfelismerő rendszer pontosabban azonosítsa a beszéd különböző elemeit. A szegmentálás automatizált módszerei közé tartozik a hangjelek energia szintjének észlelése, amely segít meghatározni, hol kezdődnek és végződnek az egyes szavak vagy mondatok.

A következő lépés a jellemzők kinyerése, amely során a hangjelből olyan adatokat nyerünk ki, amelyek leírják a beszéd akusztikus tulajdonságait. A jellemzők kinyerése kritikus fontosságú, mert ezek az adatok lesznek az alapjai a beszédfelismerési folyamatnak.

A hangfelismerésben segíthet a Fourier-transzformáció, ami egy alapvető módszer a hangjel frekvenciakomponenseinek kinyerésére. A Fourier-transzformáció lehetővé teszi, hogy a hangjelet frekvencia térbe transzformáljuk, ahol könnyebben elemezhetők a különböző frekvenciák és azok amplitúdói. Ez különösen hasznos a beszéd elemzésében, mivel a különböző hangok különböző frekvenciákat tartalmaznak [6].

Smoothie to Recipe



2.ábra: Fourier-transzformáció

<https://betterexplained.com/articles/an-interactive-guide-to-the-fourier-transform/>

A 2. ábrán láthatunk egy általános magyarázatát a Fourier-transzformációnak. Képzeljük el, hogy van egy ízletes smoothienk, amely különféle gyümölcsökből, például

eperből, banánból és narancsból készült. Bár a smoothie-ban már minden összekeveredett, a Fourier-transzformáció olyan, mintha lenne egy varázslatos eszközünk, amivel visszafejthetjük, hogy pontosan milyen gyümölcsök és milyen arányban kerültek bele. Ez a varázslat lehetővé teszi, hogy külön-külön lássuk az eper, a banán és a narancs ízét, még akkor is, ha már összekeveredtek.

Ugyanígy, a Fourier-transzformáció segítségével egy bonyolult, összetett hangjelet (a smoothie) lebontunk egyszerű, alapvető frekvenciákra (az összetevő gyümölcsökre). Így könnyebben megérthetjük, hogy egy hang miből tevődik össze, mely frekvenciák dominálnak benne, és ezek hogyan viszonyulnak egymáshoz.

Egy másik ága a beszédfelismerésnek az előre definiált hangfelismerés. Az előre definiált hangfelismerési és beszédfelismerési algoritmusok különböző formában jelenhetnek meg a szoftverfejlesztésben. Ezek a technológiák lehetővé teszik a fejlesztők számára, hogy a szoftvereket beszédvezérlésűvé tegyék, vagy automatizált hangalapú parancsokat implementáljanak. Az ilyen típusú technológiák között a Microsoft .NET keretrendszerében elérhető System.Speech osztályok is kiemelkedő szerepet töltenek be.

A hangvezérlés és beszédfelismerés technológiája az utóbbi években jelentős fejlődésen ment keresztül, és számos előnnyel jár az alkalmazásfejlesztők és a felhasználók számára. Az előnyök közé tartozik a felhasználói élmény jelentős javulása, amely lehetővé teszi a természetes és intuitív interakciót az eszközökkel és alkalmazásokkal. A hangvezérlés általában gyorsabb és kényelmesebb lehet, mint a hagyományos felületek, például billentyűzet vagy érintőképernyő használata.

Azonban hátrányokkal is rendelkezik. Ezek közé tartozik a pontatlanság, amely gyakran akkor jelentkezik, ha a rendszer nem képes pontosan értelmezni a felhasználó beszédét. A zajra való érzékenység is problémát okozhat, különösen olyan környezetekben, ahol sok háttérzaj van jelen, például nyilvános helyeken vagy zajos irodai környezetben.

Célkitűzés

Az okos otthonok koncepciója egyre inkább beépül a mindennapi életünkbe, jelentős mértékben javítva a kényelem, a biztonság és az energiahatékonyság színvonalát. A szakdolgozat célja egy komplex és integrált okos otthon rendszer tervezése és megvalósítása, amely hatékonyan képes kezelni a különböző automatizálási feladatokat és kihívásokat.

Ennek a szakdolgozatnak az elsődleges célja, hogy alaposan bemutassa az okos házak rendszereit, különös tekintettel arra az aspektusra, amely az emberi hang felismerésére és az arra adott vezérlési reakciókra fókuszál. Az okos otthon technológiák fejlődése lehetővé teszi, hogy a lakók hangparancsok segítségével irányítsák otthonuk különböző eszközeit, így növelve a kényelem és az energiahatékonyság szintjét.

A dolgozat során célom, hogy egy működőképess rendszert tervezek és implementáljak, amely képes felismerni az emberi hangot, majd megfelelően reagálni arra különböző vezérlési parancsok végrehajtásával. Ehhez a rendszert a C# programozási nyelv segítségével fogom lemodellezni, mivel a C# lehetőséget biztosít gyors és relatíve egyszerű lakásmodellek felépítésére, továbbá jól támogatja az alapértelmezett mikrofonbeállítások automatikus használatát.

Felhasznált anyagok és eszközök

C# Programozási Nyelv

A szakdolgozat során a C# programozási nyelvet használtam, amely a Microsoft által fejlesztett, objektumorientált programozási nyelv [7]. A C# előnyei közé tartozik a nyelv szintaktikai egyszerűsége, a kiterjedt standard könyvtárak és a fejlett fejlesztői eszközök támogatása. Az alábbiakban részletezem a C# nyelv használatának főbb előnyeit és alkalmazási területeit a dolgozatomban:

1. Objektumorientált Programozás (OOP):

A C# nyelv támogatja az objektumorientált programozás alapelveit, mint az öröklés, a polimorfizmus és az enkapszuláció [8]. Ezek az alapelvek lehetővé tették a rendszer moduláris felépítését, ahol az egyes komponensek külön osztályokban kerültek implementálásra, így megkönnyítve a karbantartást és a bővítést.

2. Fejlett Nyelvi Főbb Jellemzők:

A C# nyelv számos fejlett nyelvi jellemzőt kínál, mint például a LINQ (Language Integrated Query), aszinkron programozás (async/await), tulajdonságok (properties), és lambda kifejezések. Ezek a jellemzők nagymértékben hozzájárultak a kód olvashatóságához és hatékonyságához.

3. Kiterjedt Standard Könyvtárak:

A C# nyelvhez tartozó .NET keretrendszer kiterjedt standard könyvtárakat kínál, amelyek számos kész funkciót biztosítanak, például adatbáziskezeléshez, fájlműveletekhez, hálózati kommunikációhoz és grafikus felhasználói felületek létrehozásához. A dolgozatomban különösen hasznosnak bizonyultak a Windows Forms és a System.Speech könyvtárak, amelyek segítségével hangvezérlést és grafikus felületet hoztam létre. A System.Speech.Recognition könyvtárban vannak a szükséges adatbázisok és függvények, hogy egy hangot a program tudjon értelmezni és felismerni és a System.Speech.Synthesis segítségével egy általam beállított hangon meg is tudom szólaltatni, így sokkal rezponzívabb és felhasználóbarátibb demót tudok létrehozni akárkinek aki ki akarja próbálni a szimulált okos házat.

Visual Studio Fejlesztőkörnyezet

A Visual Studio egy erőteljesen integrált fejlesztői környezet (IDE), amely jelentősen megkönnyíti a teljes szoftverfejlesztési folyamatot [9]. Ez a környezet lehetővé teszi a fejlesztők számára, hogy egy helyen írják, szerkesszék, hibakeressék, építsék és telepítsék alkalmazásaikat. A Visual Studio számos programozási nyelvet támogat, beleértve a C++, C#, JavaScript, TypeScript, Python és másokat, így széleskörű projektekhez ideális eszköz.

Néhány előnye a Visual Studiónak:

1. Átfogó fejlesztői eszközök: A Visual Studio tartalmazza a kódszerkesztőket, fordítókat, hibakereső eszközöket és forráskód-kezelést. Ezek az eszközök integráltan működnek együtt, ami hatékonyabbá teszi a fejlesztést és a problémamegoldást.
2. Testre szabható telepítő: A fejlesztők kiválaszthatják, hogy mely eszközöket és funkciókat telepítik, így csak azokat az összetevőket használják, amelyekre valóban szükségük van. Ez nemcsak helyet takarít meg, hanem a rendszer erőforrásait is optimalizálja.
3. AI-támogatott fejlesztés: A Visual Studio beépített mesterséges intelligenciával segíti a kódírást és a hibák megelőzését. Az IntelliCode funkció automatikusan ajánl javaslatokat és kódrészleteket a fejlesztők munkájának megkönnyítésére.
4. Verziókezelés integráció: A beépített Git funkciók lehetővé teszik a forráskód-kezelést közvetlenül az IDE-n belül, támogatva a verziókövetést, ágazások kezelését és az együttműködést.
5. Valós idejű együttműködés: A Visual Studio Live Share funkció lehetővé teszi, hogy több fejlesztő egyszerre dolgozzon ugyanazon a projekten, valós időben osztozva a kódbázison és a hibakeresési munkameneteken.

A Visual Studio több kiadásban is elérhető, beleértve a Professional, Enterprise és a Community verziókat. Szakdolgozatomban a Community verziót használtam, amely ingyenes és nyílt forráskódú, így széles körben hozzáférhető a fejlesztők számára.

A projektem során a Windows Form technológiát alkalmaztam, mivel ebben akartam létrehozni a saját házam alaprajzát. Ez lehetővé tette számomra, hogy leszimuláljam a lakásom felokosított változatát. A Windows Form segítségével egy interaktív felületet hoztam

létre, ahol különféle okos eszközöket és automatizálási rendszereket integráltam a lakásom virtuális modelljébe. Ez a megközelítés nemcsak a programozási ismereteimet mélyítette el, hanem gyakorlati tapasztalatot is nyújtott az IoT és a házautomatizálási rendszerek terén.

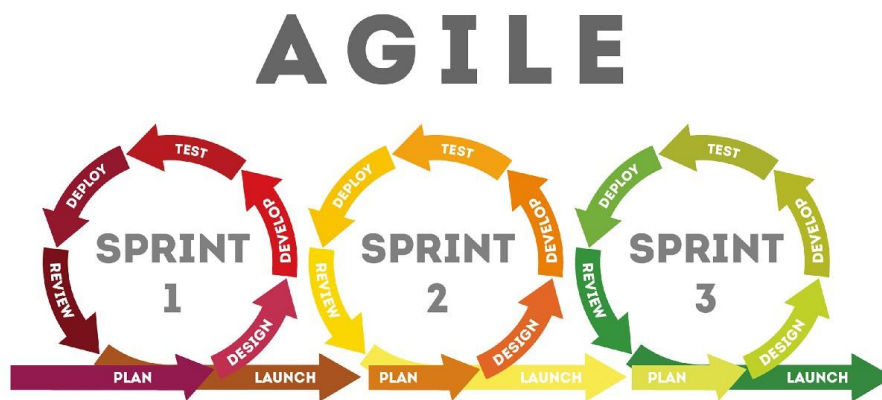
Összességében a Visual Studio fejlesztőkörnyezet hatékony eszköz a szoftverfejlesztés minden szakaszában, és számos előnnyel jár a programozók számára, akik különféle nyelveken és platformokon dolgoznak.

Alkalmazott módszerek

Agilis szoftverfejlesztés

A szakdolgozatomhoz az agilis módszertant választottam, mivel ez a megközelítés lehetővé teszi a gyors és rugalmas reakciót a felmerülő problémákra és változó követelményekre. Az agilis fejlesztési módszertan előnye, hogy rövid, iteratív ciklusokban dolgozunk, ahol minden egyes iteráció végén működőképes részegységet hozunk létre. Ez különösen fontos volt a hangfelismerés funkció fejlesztése során, mivel a rendszer gyakran nem működött megfelelően az első próbálkozások alkalmával.

A hangfelismerés egy összetett és érzékeny technológia, amely számos külső tényezőre reagálhat, mint például a háttérzaj, a beszélő hangszíne és sebessége, valamint az akcentusok. Az első iterációk során gyakran tapasztaltam, hogy a rendszer nem ismerte fel helyesen a hangparancsokat, vagy tévesen értelmezte azokat. Az agilis megközelítés lehetővé tette, hogy gyorsan azonosítsam ezeket a hibákat, és azonnal elkezdjek dolgozni a javításukon a következő iterációban.



3.ábra: Agilis szoftverfejlesztés

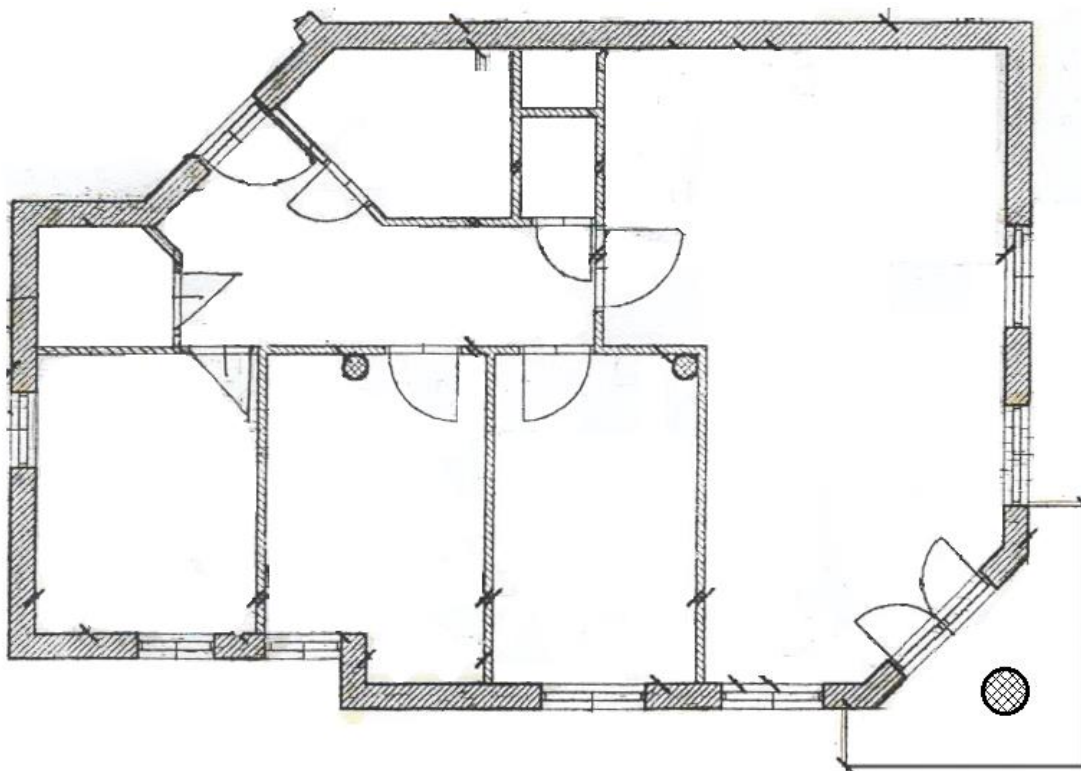
<https://pallas70.hu/kepzes-hirek/cikkek/agile-modszer-a-szoftverfejlesztesben>

A munkámat 3 sprintre osztottam, mint a 3. ábrán:

Sprint 1: Előkészítés és Alapok Lefektetése

Cél: A projekt alapjainak és infrastruktúrájának felállítása, az alapvető funkciók meghatározása és az első prototípus elkészítése.

Itt gyakorlatilag elkészítettem egy kezdetleges alaprajzot és pár grafikát, hogy mi mit reprezentáljon, itt készült el a 4.ábrán lévő alaprajz.



4.ábra: Lecsupaszított alaprajz.

Sprint 2: Funkciók Fejlesztése és Finomhangolása

Cél: A projekt főbb funkcióinak fejlesztése, kódolás és az első verzió készítése a teljes funkcionalitás biztosítása érdekében.

Itt elkészültem a lakás alaprajzával és beüzemeltem a hangfelismerő rendszert

Sprint 3: Tesztelés és Végző Finomítás

Cél: Az alkalmazás tesztelése és a visszajelzések alapján a végző finomítások és hibajavítások elvégzése a stabil és kész termék előállítása érdekében.

Itt elkészültem a légkondicionáló és a zárrendszerrel.

System.Speech

A szakdolgozatomban a hangvezérlés és -felismerés megvalósításához a .NET keretrendszer részeként elérhető System.Speech[10] névtérben található SpeechRecognitionEngine osztályt használtam. Ez az osztály széleskörű és rendkívül hasznos

funkciókkal van ellátva, amelyek lehetővé teszik a beszéd alapú parancsok felismerését és kezelését a C# nyelven írt alkalmazásokban. A `SpeechRecognitionEngine` osztály használata egyszerűvé és hatékonyá teszi a különféle beszédvezérlési feladatok megvalósítását, mivel beépített módszerekkel rendelkezik a hangok elemzésére, a felismert szavak azonosítására és a megfelelő válaszok generálására. Ez a funkcionalitás különösen hasznos olyan alkalmazások fejlesztésében, amelyek interaktív hangvezérlést igényelnek, és lehetőséget ad a felhasználói élmény jelentős javítására.

Problémák implementálásnál

A hangfelismerés implementálása során számos kihívással szembesültem, különösen a mikrofon érzékelésével kapcsolatban. Kezdetben a rendszer nem érzékelte a mikrofont, és ennek okát megtalálni időbe tellett. Az ilyen problémák gyakori okai közé tartozhatnak a hardverhibák, a rosszul beállított vagy hiányzó illesztőprogramok, illetve a Windows operációs rendszer általános hangbeállítási problémái.

Az első lépés a hibakeresés során a hardver ellenőrzése és a mikrofon csatlakozásának megfelelő működésének biztosítása volt. Ellenőriztem a mikrofon kábeleit, illetve más eszközökhöz való csatlakoztatását, hogy kizárjam a fizikai hibák lehetőségét.

A második lépés az illesztőprogramok ellenőrzése és frissítése volt. A régi vagy hiányzó illesztőprogramok gyakran okozhatnak problémákat a hangfelismerés során, ezért fontos volt biztosítani, hogy a mikrofonhoz tartozó illesztőprogramok a legfrissebb verzióban legyenek telepítve, tehát hogy gyors és egyszerű legyen mivel AMD-t használok ezért lehetőségem volt a Radeon szoftvert használni, ezzel automatikusan frissítettem az összes illesztőprogramot.

A harmadik lépés az operációs rendszer hangbeállításainak ellenőrzése és finomhangolása volt. Néhány alkalommal az operációs rendszer hangbeállításai nem voltak megfelelően konfigurálva, ami negatívan befolyásolták a hangfelismerést. Ezek közé tartozott például a mikrofon hangerejének beállítása vagy az alapértelmezett mikrofon kiválasztása.

A fenti hibajavítási kísérletek bár néhány esetben hasznosak voltak, sokáig tartott megállapítani, hogy a probléma gyökere abban rejlett, hogy a Windows operációs rendszer nyelve nem volt angolra állítva. Ennek eredményeként a `Speech.Recognition` motor nem tudta megfelelően kezelni az alapértelmezett audiót, és így nem érzékelte a mikrofont. Ennek a felismerése után az operációs rendszer nyelvének módosítása és az alapértelmezett hangbeállítások újbóli konfigurálása megszüntette a problémát, és a hangfelismerési funkció

megfelelően működött. Ez a tapasztalat rámutat arra, hogy az operációs rendszer alapvető beállításai is jelentős hatással lehetnek az alkalmazások működésére, és fontos lehet ezeknek a beállításoknak a megfelelő konfigurálása a problémák megelőzése érdekében.

Eredmények

Fonetikus írás

Sajnos a fejlesztés során viszonylag hamar rá kellett döbbernem, hogy a Microsoft által biztosított hangfelismerő rendszer a magyar nyelvet nem támogatja, van külön paramétere, ahol meg lehet adni a felolvasó és értelmező nemét, életkorát és nemzetiségét is, de sajnos ebből magyart nem tudja értelmezni. Ezt a problémát fel lehet oldani, fonetikusán leírunk minden hangot ahogy egy angol nyelven beszélő olvasna egy magyar szöveget. Ehez segítségemre volt az IPA (International Phonetic Alphabet) [11] nemzetközi fonetikai ábécé. Ez gyakorlatilag egy táblázat az összes magyar hangot átírja egy nemzetközileg is elfogadott egységes szimbólummá. Tehát a feladatom annyi volt, hogy az általam felhasznált hangokat kikeressem az angol IPA táblázatból. Ennek eredménye az 1. táblázat lett.

1. táblázat: Magyar Angol IPA táblázat

Magyar Hang	IPA Jelölés	Fonetikusan (Angol Betűkészlettel)
a	/ɒ/	o
á	/a:/	ah
b	/b/	b
c	/ts/	ts
cs	/tʃ/	ch
d	/d/	d
e	/ɛ/	eh
é	/e:/	ay
f	/f/	f
g	/g/	g
gy	/j/	dy
h	/h/	h
i	/i/	ee
j	/j/	y
k	/k/	k
l	/l/	l
ly	/j/	y
m	/m/	m
n	/n/	n
ny	/ɲ/	ne
o	/o/	aw
ó	/o:/	oh
ö	/ø/	ö
ő	/ø:/	ő
p	/p/	p

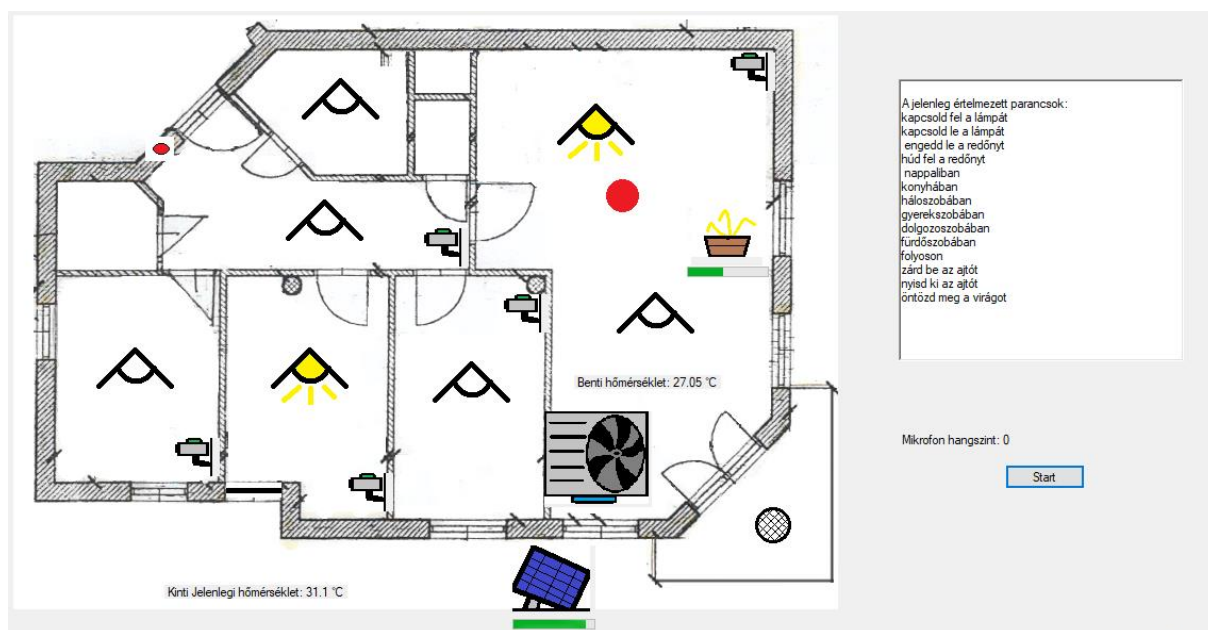
r	/r/	r
s	/ʃ/	sh
sz	/s/	s
t	/t/	t
ty	/c/	ty
u	/u/	oo
ú	/u:/	oo
ü	/y/	ü
ű	/y:/	ű
v	/v/	v
z	/z/	z
zs	/ʒ/	zh

A táblázatban szereplő angol megfelelők segítenek abban, hogy az angolul beszélők hogyan ejtenék ki a magyar hangokat, így a program könnyebben azonosítani tudja azokat. A táblázat áttekintést nyújt arról, hogy mely angol hangok milyen módon feleltethetők meg a magyar hangoknak, ami segíti a hangfelismerést és a felhasználóval való kommunikációt. Lényegében ezek a hangok lettek behelyettesítve mindenhol, ahol a nyelvek közti különbség miatt problémába ütközhetünk. Ennek oka részben az, hogy az angol hangok és azok fonetikus jelölése nem minden esetben fedik le a magyar hangok sokféleségét és varianciáját. Például a magyar "tyúk" szó hangjainak pontos ábrázolása angol fonetikával nehézségekbe ütközik, mivel az angol nyelvben nincs megfelelő hang a "ty" hang számára. Emellett a magyar hangokat tartalmazó szavak angolosítása is problémákat okozhat a hangfelismerés szempontjából. Mivel az angol beszélők nem feltétlenül ismerik vagy használják a magyar kiejtést, az általuk használt angolosított verziók eltérhetnek a helyes magyar kiejtéstől, ami tovább nehezítheti a pontos felismerést. Ezt rendszerint szavak, szókapcsolatok összerakásával és szétszedésével próbáltam kiküszöbölni, például a „Szia üdvözöllek az okos házban” felolvasásból „Seeya, oodvehzohlekaz ókosh hazban”. Ebben a példában az is látszik, hogy különbség van az elől- és hátul-képzetes hangok között bontsuk le ezt a példát. A "Szia" helyett "Seeya" formátumot használtam, hogy az angol beszélők számára könnyebben kiejthető legyen. Az "üdvözöllek" helyett "oodvehzohlekaz" módon írtam le, mivel az "ü" hang angolul nincs jelen, és a "v" és "z" hangokat angolosítva írtam, ezen felül a végére odakerült a következő szó az „az” mivel a hangfelismerő nem tudott mit kezdeni ezzel a két betűs szóval és rendszeresen csak betűzte. Az "okos házban" pedig "ókosh hazban" formában szerepel, mivel az "okos" szóban az "o" hang elől, míg a "házban" szóban a "h" hang hátul képződik. Ezáltal az angol hangfelismerő rendszer pontosabban tudja azonosítani és kezelni ezeket a szavakat.

Ezek az eltérések és nehézségek magyarázatot adhatnak arra, hogy miért nem mindig tökéletes a hangfelismerés, még akkor sem, ha a felhasználó megpróbálja fonetikusán ábrázolni a magyar szavakat az angol hangrendszerben. A pontos és megbízható hangfelismerés elérése érdekében további finomhangolásokra és fejlesztésekre lehet szükség a rendszerben.

Hangvezérelt Virtuális Otthon

Az alkalmazásom lehetővé teszi a felhasználók számára, hogy hangutasítások segítségével irányítsanak egy virtuális otthon eszközeit és funkcióit.



5.ábra: Hangvezérelt virtuális otthon

Az alkalmazás a következő főbb jellemzőkkel rendelkezik:

Hangfelismerés és Hangszintérzékelés: Az alkalmazás felismeri a felhasználó által kimondott hangutasításokat, és megfelelően reagál rájuk. Emellett képes megjeleníteni a mikrofon által érzékelt hangszintet.

Virtuális Környezet: A program egy virtuális környezetet biztosít, amelyben különböző helyszínek és azokban elhelyezett eszközök találhatók. A felhasználó képes navigálni a környezetben és vezérelni az eszközöket hangutasítások segítségével.

Interaktív Funkciók: Az alkalmazás interaktív módon lehetővé teszi a felhasználók számára, hogy változtassanak a virtuális környezetben található helyszínek között, valamint vezéreljék az ott található eszközöket.

Használati útmutató:

A virtuális okosház programot, ha elindítjuk kezdetben a házon kívül találjuk magunkat

Program Indítása:

Amikor elindítod a programot, a kezdőképernyőn megjelenik a ház alaprajza.

A tartózkodási helyedet egy piros pont jelöli, amely a házon kívül helyezkedik el.

Hangvezérlés Aktiválása:

A program indítása után kattints a „Start” gombra a hangvezérlés aktiválásához.

A hangvezérlés aktiválásakor a rendszer üdvözlő üzenetet mond, és készen áll a parancsok fogadására.

Ajtó Kinyitása:

A házba való belépéshez először nyisd ki az ajtót hangvezérléssel. Mondd például: „nyisd ki az ajtót”.

Ha az ajtó nyitva van, a piros pont beléphet a házba.

Mozgás a Házban:

Az ajtó kinyitása után mondd ki annak a szobának a nevét, ahová be szeretnél lépni. Például: „menj a nappaliba”.

A piros pont az adott szobába mozog, és megjelenik a szoba alaprajza.

Szobák és Funkciók

Nappali:

A nappaliban a következő parancsokat használhatod:

- „kapcsold fel a lámpát” – a lámpa felkapcsolása
- „kapcsold le a lámpát” – a lámpa lekapcsolása.
- „engedd le a redőnyt” – a redőny leengedése.

- „húzd fel a redőnyt” – a redőny felhúzása.
- „öntözd meg a virágot” – a virág megöntözése.

Konyha:

A konyhában a következő parancsokat használhatod:

- „kapcsold fel a lámpát” – a lámpa felkapcsolása.
- „kapcsold le a lámpát” – a lámpa lekapcsolása.
- „engedd le a redőnyt” – a redőny leengedése.
- „hozd fel a redőnyt” – a redőny felhúzása.
- „öntözd meg a virágot” – a virág megöntözése.

- Hálószoza:

A hálószozában a következő parancsokat használhatod:

- „kapcsold fel a lámpát” – a lámpa felkapcsolása.
- „kapcsold le a lámpát” – a lámpa lekapcsolása.
- „engedd le a redőnyt” – a redőny leengedése.
- „hozd fel a redőnyt” – a redőny felhúzása.

Gyerekszoza:

A gyerekszozában a következő parancsokat használhatod:

- „kapcsold fel a lámpát” – a lámpa felkapcsolása.
- „kapcsold le a lámpát” – a lámpa lekapcsolása.
- „engedd le a redőnyt” – a redőny leengedése.
- „hozd fel a redőnyt” – a redőny felhúzása.

Dolgozószoza:

A dolgozószozában a következő parancsokat használhatod:

- „kapcsold fel a lámpát” – a lámpa felkapcsolása.

- „kapcsold le a lámpát” – a lámpa lekapcsolása.
- „engedd le a redőnyt” – a redőny leengedése.
- „hozd fel a redőnyt” – a redőny felhúzása.

Fürdőszoba:

A fürdőszobában a következő parancsokat használhatod:

- „kapcsold fel a lámpát” – a lámpa felkapcsolása.
- „kapcsold le a lámpát” – a lámpa lekapcsolása.

Folyosó:

A folyosón a következő parancsokat használhatod:

- „kapcsold fel a lámpát” – a lámpa felkapcsolása.
- „kapcsold le a lámpát” – a lámpa lekapcsolása.

Kilépés a Házból

Ha ki szeretnél lépni a házból, mondd a „kint” parancsot.

A piros pont kilép a házból, és ismét a házban kívül található.

Ajtó Bezárása:

Ha befejeztél a ház irányítását, ne felejtse el bezárni az ajtót. Mondd például: „zárd be az ajtót”.

Az ajtó bezáródik, és a piros pont nem tud többé bemenni, amíg újra ki nem nyitod.

Ezen felül bármikor kérhetsz segítséget és az összes parancsot kilistázza a program.

Ez az útmutató remélhetőleg segít eligazodni az okosház program használatában.

Automatikus rendszerek:

A lakás automatikusan folyamatosan próbálja 24°C-ra állítani a benti hőmérsékletet, ehhez automatikusan vizsgálja a jelenlegi benti hőmérsékletet és automatikusan eldönti, hogy fűteni vagy hűtenie kell, ezen felül van egy napelem is, ami segít a ház energiaellátását fenntartani.

További tervek

További tervek között szerepel a magyar nyelvű felismerés pontosítása és javítása. Azon dolgozom, hogy az alkalmazás még jobban megértse és kezelje a magyar nyelv sajátosságait és kifejezéseit. Emellett egy automatikus rendszert fejleszték, amely képes lesz egy alaprajz alapján felismerni és megkülönböztetni a különböző szobákat. A felhasználók számára lehetőséget biztosítok majd arra, hogy saját preferenciáiknak megfelelően nevezzék el ezeket a helyiségeket. Továbbá tervezem, hogy a légkondicionálót is hangvezérléssel lehessen irányítani, így a felhasználók könnyedén beállíthatják és szabályozhatják a kívánt hőmérsékletet a különböző helyiségekben. Ezek az új funkciók és fejlesztések hozzájárulnak az alkalmazás funkcionalitásának és felhasználói élményének további javításához.

Összefoglalás

Az Intelligens Hangvezérlésű Otthoni Automatizációs Rendszer egy innovatív megoldás, amely lehetővé teszi a felhasználók számára, hogy hangutasításokkal irányítsák otthonuk különböző eszközeit és funkcióit. Az alkalmazás számos intelligens funkciót kínál, beleértve a világítás, a fűtés, a légkondicionálás és az ablakok vezérlését.

Az alkalmazás elsődleges célja az otthoni kényelem és hatékonyság növelése. A felhasználók könnyedén vezérelhetik otthonuk különböző eszközeit és funkcióit hangutasításokkal anélkül, hogy megérintenék a kapcsolókat vagy érintőképernyőket. Ezáltal időt és energiát takaríthatnak meg, miközben kényelmesebbé teszik mindennapi életüket ez különösen akkor lehet előnyös, ha valamilyen testi visszamaradottsággal él a felhasználó akkor ezeket az egyszerű kényelmi funkciókat tudja úgy használni, hogy gyorsabb legyen és kevesebb problémát okozzanak ezek az otthoni tevékenységek. Vannak automatikus rendszerek is amelyek segítik az okos otthon élhetőbbé tételét. Interfészek és képernyők vagy kapcsolók nélkül irányíthatjuk a lakásunkat, ami hozzáad a kényelemhez és egyszerű felhasználáshoz.

A rendszer különlegessége, hogy magyar nyelven lehet vele kommunikálni. Ez az újítás a magyar nyelv és kultúra mérete és elterjedtsége miatt problémákat okozhat a jövőben amíg a hangfelismeréshez kevés hang adatunk van. Ha a piacon megjelenik az igény, hogy magyar nyelven tudjunk tesztelni akkor elérhetjük, hogy a Visual Studio is képes legyen egyszerűen és akadálymentesen feldolgozni ezeket az igényeket.

Minden esetre amíg nincs hivatalom Microsoft update azzal kapcsolatban, hogy a magyar beszédet is felismerjék és megszólaltassák, addig remélem az összeszedett magyar-angol fonetikus hangzástáblázatom segít azoknak, akik hasonló témával akarnak foglalkozni.

Remélem jelenlegi szakdolgozatommal sikerült inspirálnom az olvasót arra, hogy elgondolkozzon az otthonába beüzemelhető okos eszközök használatán. Ez a rendszer napról napra fejlődik, és különböző Android és Ios rendszerek bizonyos aspektusban előrébb tartanak, mint ami ennek a dolgozatnak a célja volt viszont amit keveseknek sikerült elérni az az, hogy egy rendszer elérhető áron és megbízható biztonsági intézkedésekkel tud biztosítani egy egységes átfogó modern élményt.

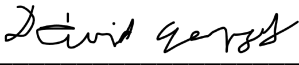
Irodalomjegyzék

1. Általános leírás az okos házról: https://hu.wikipedia.org/wiki/Intelligens_épület
2. Légkondi: <https://www.digitaltrends.com/home/air-conditioner-deals/>
3. Kamera rendszerek: <https://www.security.org/security-cameras/best/>
4. IoT(coursera):https://www.coursera.org/articles/internet-of-things?adpostion=&utm_medium=sem&utm_source=gg&utm_campaign=B2C_EMEA_coursera_FTCOF_career-academy_pmax-multiple-audiences-country-multi-set2&campaignid=20882109092&adgroupid=&device=c&keyword=&matchtype=&network=x&devicemodel=&adposition=&creativeid=&hide_mobile_promo&gad_source=1&gclid=CjwKCAjwupGyBhBBEiwA0UcqaPlrh-ZiXxwyxvdVuohbw0fv1NIHPGCrTqz-sd-dJJZ2w9nsEQnGuhoCx9cQAvD_BwE
5. Hangfelismerés:https://docs.google.com/document/d/1T8ZhS8BZc_83t3V3EAe_8M4J7aReQz3CWvzD3V57G_Tg/edit,
https://en.wikipedia.org/wiki/Windows_Speech_Recognition
6. Fourier-Transzformáció:<https://betterexplained.com/articles/an-interactive-guide-to-the-fourier-transform/>
7. C#: [https://en.wikipedia.org/wiki/C_Sharp_\(programming_language\)](https://en.wikipedia.org/wiki/C_Sharp_(programming_language))
8. Objektumorientált programozás: https://en.wikipedia.org/wiki/Object-oriented_programming
9. Visual Studio: <https://learn.microsoft.com/en-us/visualstudio/get-started/visual-studio-ide?view=vs-2022>
10. System.Speech:<https://learn.microsoft.com/en-us/dotnet/api/system.speech.synthesis?view=net-8.0>
11. Fonetikai ábécé https://hu.wikipedia.org/wiki/Nemzetközi_fonetikai_ábécé

Nyilatkozat

Alulírott, Dávid Gergely, Programtervező informatikus szakos hallgató, kijelentem, hogy a szakdolgozatban ismertetettek saját munkám eredményei, és minden felhasznált, nem saját munkából származó eredmény esetén hivatkozással jelöltem annak forrását.

Dátum (Szeged, 2024, 05, 16)



aláírás

Mellékletek

A dokumentum mellett van egy videó a beszédfelismerés működéséről