Recognition of Ultrasonic Multi-Echo Sequences for Autonomous Symbolic Indoor Tracking

André Stuhlsatz
Collaborative Virtual and Augmented Environments Department
Fraunhofer Institute for Applied Information Technology FIT
Schloss Birlinghoven
53754 Sankt Augustin, Germany
Email: andre.stuhlsatz@fit.fraunhofer.de

Abstract

This paper presents an autonomous symbolic indoor tracking system for ubiquitous computing applications. The proposed approach is based upon the assumption that topologically discriminable information can be assigned explicitly to different spaces of a given indoor environment. On that assumption, continuous Time-of-Flight (ToF) measurements of echo-bursts obtained from four orthogonally and coplanarly mounted ultrasonic transducer are used to learn a stochastic room model. While the individual acoustic representation of space is captured using Gaussian mixture densities, the stochastic variabilities in the moving direction of a person are modeled by Hidden-Markov-Models (HMMs). Experiments within a six room environment resulted in a room recognition rate of 92.21% and a room sequence recogntion rate of 66.00% without any pre-fixed devices.

1. Introduction

Although many high accuracy positioning systems exist, indoor location-awareness remains a challenging task. For example, the Global Positioning System (GPS) provides very accurate positioning information but suffers from poor indoor applicability because of signal attenuation inside buildings. In addition, indoor environments usually consist of metal and other materials and these negatively affect the estimation of position permitting by RF-signals or inertial navigation system (INS). Therefore, alternatives based upon different physical principles have been developed in recent years to overcome such difficulties [1, 2, 3, 4, 5]. All of these approaches make use of numerousness installed sensors at precisely known locations. Accuracies in position of a few centimeters are attainable with a high degree of logistic effort and expense for an area-wide installation of sensors. In many ubiquitous computing applications, a highly accurate position is not necessary. Instead, a rough resolution of a few square meters associated with labeled places is sufficient to permit the monitoring of moving persons or some equipment within a building, therefore it is also called *symbolic localization*. For example, imagine a blind person who will be guided from room to room through a hospital or to a bookshelf inside a public library. The requirements for symbolic localization are of a quite different nature, namely symbolic localization has to be self-contained, inexpensive and easy adaptable to changing surroundings. This paper addresses all these issues and a novel symbolic localization framework will be presented.

2. Symbolic Localization Framework

The proposed framework relies on the assumption that different rooms or parts of rooms (*subrooms*) are distinguishable by means of their unique topological characteristics (Fig. 1). This means as long as enough static structure

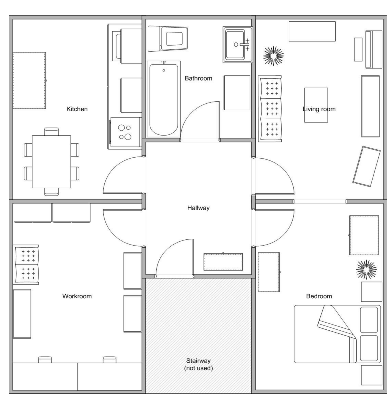

Figure 1. Test environment.

Figure 2. Ultrasonic transceiver module SRF08 from Devantech [7].

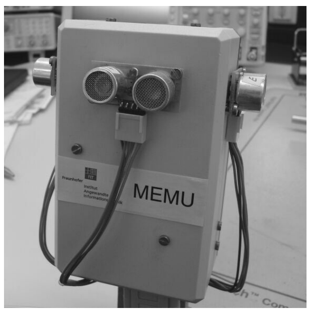

Figure 3. Measurement unit consisting of four orthogonally and coplanarly mounted SRF08.

is provided in terms of furniture, doors, windows and the outline of the rooms, it is possible to use range measurements (*topological features*) for a classification into symbolic representatives, like KITCHEN or BATHROOM. The proposed method is divided in two building blocks: a hardware component for ultrasonic range measurements and a stochastic room model for continuous classification of features in (sub-)rooms.

2.1. Hardware implementation

Because a ranging resolution of less than a few centimeters is not required but instead a low price is an important objective, simple commercial available ultrasonic transceivers are used (Fig. 2). The SRF08 transceiver module features a range measurement of 0.03-6m, a programmable gain control and a low current drain for mobile applications of 15mA typical/3mA standby. The module emits a 40kHz ultrasonic burst and is able to detect up to 17 echoes in series within a 36ms acquisition time. This feature is crucial, because it allows it to construct a depth image of space. To acquire as many of the topological features of the environment as possible, four SRF08 modules are assembled orthogonally and coplanarly to a multi-echo measurement unit (*MEMU*) (Fig. 3). Inside the MEMU, a PIC microcontroller initiates sequentially firing of every module's sonar burst, performs the pre-processing

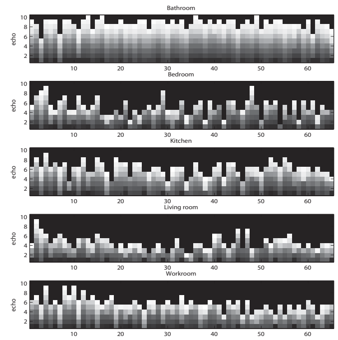

Figure 4. This figure shows topological features of five rooms of the test environment measured with the MEMU.

Figure 5. The SYLOC prototype consisting of the MEMU and a notebook fixed on a wearable baby carrier.

and handles the communication to a mobile computer for post-processing. A full cycle completes in about 440ms resulting in a feature rate of about nine features per second. In (Fig. 4) some samples are shown, at which the y-axis relates to the echo number, the x-axis relates to the feature vector number (respectively time) and the color relates to the distance (long distances are brighter than short distances). Note, measurements of the bathroom above the fifth echo contain little distinguishable information. This is because of the very small outline of the bathroom and the resulting multi-echo interferences. The MEMU together with a laptop were fixed on a wearable baby carrier (Fig. 5) and served as experimental prototype of the proposed self-contained symbolic localization system (*SYLOC*).

2.2. Stochastic Modeling

The feature generation observed by the SYLOC-system during a walk through different rooms has been modeled

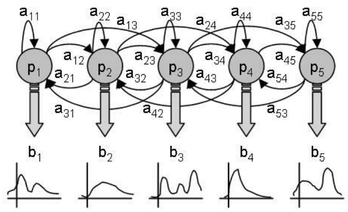

Figure 6. A continuous Hidden-Markov-Model (HMM).

under the assumption of two dependent stochastic processes. The first process generates a finite discrete random walk $(\omega_1, \ldots, \omega_T)$ through a discretized environment $\Omega := \{p_1, \ldots, p_K\}$ along randomly chosen positions $\omega_t \in \Omega$ at discrete times $t \in \mathcal{T} := \{1, \ldots, T\}$. Different speeds of the walker are modeled by allowing to skip states (faster) or to repeat states (slower). The second process produces randomly distributed echo measurements $\boldsymbol{x} \in I\!R^d$ dependent on the actual position ω_t as well as on the different reflections at static elements in space and interferences of the ultrasonic waves. Although, the actual path a user takes through the environment is not known explicitly, it is possible to estimate the most probable room or subroom from a sequence of observed measurements during a random walk applying *Hidden-Markov-Models* (HMMs) (Fig. 6). Hidden-Markov-Models are well-known, and have been successfully used in speech recognition since many years [8]. In general, a stochastic process under the assumption that the probability to be in position ω_t depends only on the predecessor position ω_{t-1}, i.e.

$$P(\omega_t = p_{i_t} | \omega_{t-1} = p_{i_{t-1}}, \omega_{t-2} = p_{i_{t-2}} \ldots) = \quad (1)$$
$$P(\omega_t = p_{i_t} | \omega_{t-1} = p_{i_{t-1}}) \; \forall t \in \mathcal{T} : 1 \leq i_t \leq K, \quad (2)$$

is called a *first-order Markov process*. Independence of the absolute time and writing $k := i_t, l := i_{t+1}$ enables it to define the *transition probabilities* from the current position to the next position as

$$a_{kl} := P(\omega_{t+1} = p_l | \omega_t = p_k) \; \forall 1 \leq k, l \leq K \quad (3)$$

which have the properties

$$a_{kl} \geq 0 \; \forall 1 \leq k, l \leq K \quad (4)$$
$$\sum_{l=1}^{K} a_{kl} = 1 \; \forall 1 \leq k \leq K. \quad (5)$$

Under the condition that the actual position (or state of the HMM) ω_t equals p_k, and the measurement error caused by different reflections and interferences at objects in the surrounding is well modeled using *Gaussian mixture densities*

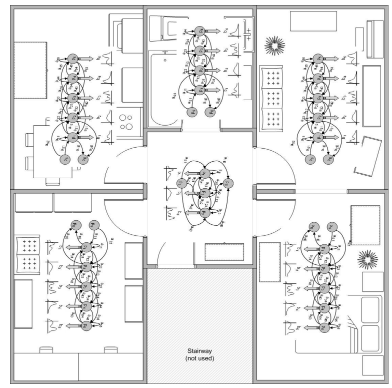

Figure 7. A HMM is assigned to each location.

consisting of M linear combined Gaussians, the observation density of a measurement $\boldsymbol{x}$ is given by

$$b_k(\boldsymbol{x}) := p(\boldsymbol{X} = \boldsymbol{x} | \omega_t = p_k, \boldsymbol{\lambda}^k) = \quad (6)$$
$$\sum_{m=1}^{M} \frac{\alpha_m^k}{\sqrt{(2\pi)^d \det(\boldsymbol{\Sigma}_m^k)}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_m^k)^{Tr}(\boldsymbol{\Sigma}_m^k)^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_m^k)}, \quad (7)$$
$$\alpha_m^k \geq 0, \sum_{m=1}^{M} \alpha_m^k = 1 \; \forall 1 \leq k \leq K, \quad (8)$$
$$\boldsymbol{\lambda}^k := (\alpha_1^k, \ldots, \alpha_M^k, \boldsymbol{\mu}_1^k, \ldots, \boldsymbol{\mu}_M^k, \boldsymbol{\Sigma}_1^k, \ldots, \boldsymbol{\Sigma}_M^k). \quad (9)$$

A HMM is fully characterized by a transition matrix $\mathbf{A} := (a_{kl})_{1 \leq k, l \leq K}$, the observation densities $\boldsymbol{b} := (b_k)_{1 \leq k \leq K}$ and the initial state distribution $\boldsymbol{\pi} := (\pi_k)_{1 \leq k \leq K}$ with

$$\pi_k := P(\omega_1 = p_k) \; \forall 1 \leq k \leq K. \quad (10)$$

Before HMMs can be applied for positioning in the SYLOC-system, three crucial questions have to be answered. In the following, these questions are stated and well-known results are given for a brief introduction in HMMs with references for further reading:

2.2.1 The recognition of rooms from observations

Given an observation sequence $\mathcal{O} := (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_T)$ of independent multi-echo range measurements obtained from the MEMU, one wants to classify the sequence $\mathcal{O}$ in one of the defined location symbols $\mathcal{L} := \{s_1, \ldots, s_L\}$, as e.g. KITCHEN, NEAR_COMPUTER or 3RD_FLOOR. Thus, from a statistical point of view one is faced with the problem of finding the most probable HMM $\boldsymbol{\gamma}^q := (\mathbf{A}^q, \boldsymbol{b}^q, \boldsymbol{\pi}^q),\; 1 \leq q \leq L$ representing a labeled location $s_q \in \mathcal{L}$ given the observation sequence $\mathcal{O}$ (Fig. 7) , i.e.

$$q^* := \underset{q \in \{1,\ldots,L\}}{\operatorname{argmax}} \frac{P(\mathcal{O}|\boldsymbol{\gamma}^q) P(\boldsymbol{\gamma}^q)}{P(\mathcal{O})}. \quad (11)$$

Note in (Fig. 7), there are particular states, one start state p_S and one end state p_E, which represent the doors of each room and are used to interconnect the HMMs for room sequence recognition given an observation sequence $\mathcal{O}$. For the optimization in (11) it is sufficient to find an optimal q^* for which the nominator $P(\mathcal{O}|\boldsymbol{\gamma}^q)P(\boldsymbol{\gamma}^q)$ is maximized. The prior $P(\boldsymbol{\gamma}^q)$ is usually defined based on particular knowledge of the problem, whereas the observation *likelihood* $P(\mathcal{O}|\boldsymbol{\gamma}^q)$ has to be estimated. Suppose an observation sequence $\mathcal{O}$, then the likelihood for the sequence given a model $\boldsymbol{\gamma}^q$ equals

$$P(\mathcal{O}|\boldsymbol{\gamma}^q) = \sum_{\mathcal{W}\in\{1,\ldots,K^q\}^T} P(\mathcal{O}|\mathcal{W},\boldsymbol{\gamma}^q)P(\mathcal{W}|\boldsymbol{\gamma}^q) \quad (12)$$

$$= \sum_{\{i_1,\ldots,i_T\}\in\{1,\ldots,K^q\}^T} \pi_{i_1} b_{i_1}(\boldsymbol{x}_1)\prod_{t=2}^{T} a_{i_{t-1}i_t} b_{i_t}(\boldsymbol{x}_t).$$

Obviously, the computation of $P(\mathcal{O}|\boldsymbol{\gamma}^q)$ is intractable in practice, even for a small number K^q of states, because of the high number of possible hidden walks $\mathcal{W}$. Fortunately, there exists a calculation procedure called the *Forward-Backward algorithm* [8] which solves this problem efficiently.

2.2.2 The optimal state sequence

As the actual path is hidden, the problem is to find the most probable path $\mathcal{W}^*$ through the environment. This can be solved by using the well-known *Viterbi-algorithm* [8]. The Viterbi-algorithm efficiently maximizes the probability of a single path $\mathcal{W}$ given the observations sequence $\mathcal{O}$ and the model $\boldsymbol{\gamma}^q$, i.e.

$$P(\mathcal{W}^*|\mathcal{O},\boldsymbol{\gamma}^q) = \max_{\mathcal{W}\in\{1,\ldots,K^q\}^T} \frac{P(\mathcal{O},\mathcal{W}|\boldsymbol{\gamma}^q)}{P(\mathcal{O})}. \quad (13)$$

First, define the probability of a partial path $(p_{i_1},\ldots,p_{i_{t-1}},p_k)$ accounting for the observations $\{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_t\}$ whereas $\omega_t = p_k$ at time t, i.e.

$$\delta_t(k) := \max_{i_1,\ldots,i_{t-1}} P(\omega_t = p_k, p_{i_1},\ldots,p_{i_{t-1}},\boldsymbol{x}_1,\ldots,\boldsymbol{x}_t|\gamma^q). \quad (14)$$

Then the most probable state sequence is found by applying the recursion

$$\delta_t(l) = b_l(\boldsymbol{x}_t)\max_k \delta_{t-1}(k)a_{kl},\ 2\le t\le T \quad (15)$$

and backtracking from the final state ω_T to the first state ω_1 traversing transitions which led to the maximal partial path probability. In (Fig. 8) a transition graph for decoding is shown, at which the vertical represents the states of the HMM and the horizontal represents the time a feature $\boldsymbol{x}_t$ is observed. The nodes represents the density $b_k(\boldsymbol{x}_t)$ and the arcs represent the transition probabilities a_{kl}. The probability of a path given the observations is simply the product of all transition and observation densities along that path.

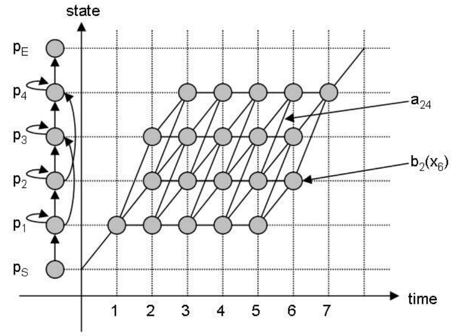

Figure 8. The Viterbi algorithm finds the most probable path through a graph.

2.2.3 Training of room-models

The training (or estimation of parameters) of a HMM for every location in $\mathcal{L}$ is by far the most difficult problem. Because the equations required to solve for an optimal set of parameters $(\mathbf{A}^*,\boldsymbol{b}^*,\boldsymbol{\pi}^*)$ are highly coupled, it is not possible to derive a solution analytically. However, an iterative technique to find a set of parameters corresponding to a local maximum of $P(\mathcal{O}|\boldsymbol{\gamma}^q)$ exists, namely the *Baum-Welch algorithm* [8]. Defining the probability of being at position p_k at time t, and position p_l at time $t+1$, conditioned on the observations $\mathcal{O}$ and the model $\boldsymbol{\gamma}^q$, i.e.

$$\xi_t(k,l) := P(\omega_t = p_k, \omega_{t+1} = p_l|\mathcal{O},\boldsymbol{\gamma}^q), \quad (16)$$

the Baum-Welch reestimation formulas for $(\boldsymbol{\pi},\mathbf{A})$ read as

$$\bar{\pi}_k := \sum_{l=1}^{K}\xi_1(k,l), \quad (17)$$

$$\bar{a}_{kl} := \frac{\sum_{t=1}^{T-1}\xi_t(k,l)}{\sum_{t=1}^{T-1}\sum_{l=1}^{K}\xi_t(k,l)}\ \forall 1\le k,l\le K. \quad (18)$$

The reestimation formulas for the parameters $\boldsymbol{\lambda}^k$ of every Gaussian mixture densities b_k are given by:

$$\bar{\alpha}_m^k := \frac{\sum_{t=1}^{T}\zeta_t(k,m)}{\sum_{t=1}^{T}\sum_{m=1}^{M}\zeta_t(k,m)}, \quad (19)$$

$$\bar{\boldsymbol{\mu}}_m^k := \frac{\sum_{t=1}^{T}\zeta_t(k,m)\boldsymbol{x}_t}{\sum_{t=1}^{T}\zeta_t(k,m)}, \quad (20)$$

$$\bar{\boldsymbol{\Sigma}}_m^k := \frac{\sum_{t=1}^{T}\zeta_t(k,m)(\boldsymbol{x}_t-\bar{\boldsymbol{\mu}}_m^k)(\boldsymbol{x}_t-\bar{\boldsymbol{\mu}}_m^k)^{Tr}}{\sum_{t=1}^{T}\zeta_t(k,m)} \quad (21)$$
$$\forall 1\le k\le K, 1\le m\le M,$$

where $\zeta_t(k,m) := P(\omega_t = p_k, \upsilon_t = m|\mathcal{O},\boldsymbol{\gamma}^q)$ is the probability to be in state p_k at time t and the mth-mixture component accounts for observation $\boldsymbol{x}_t$. Note, that both

$\xi_t(k,l)$ and $\zeta_t(k,m)$ can be calculated using the Forward-Backward algorithm. The Baum-Welch reestimation has been proven to improve the likelihood of the observation sequence, i.e. $P(\mathcal{O}|\bar{\gamma}^q) > P(\mathcal{O}|\gamma^q)$, or $\bar{\gamma}^q = \gamma^q$ [8]. Thus, starting with an initial guess γ_0^q one can iteratively produce a sequence of parameters γ_i^q with $P(\mathcal{O}|\gamma_{i+1}^q) > P(\mathcal{O}|\gamma_i^q)$ using the update formulas and terminate e.g. when the improvement is below a given threshold. Although, this procedure yields an local extremum, it is not guaranteed to find a global optimum.

2.3. Recognition of Room Sequences

Consider N randomly visited (sub-)rooms $\Gamma_N := (\gamma^{i_1}, \ldots, \gamma^{i_N}), 1 \leq i_j \leq L$ represented by trained HMMs for symbolic localization (like a sequence of continuously spoken words in continuous speech recognition). Recognizing the most probable room sequence Γ_N^* is accomplished by interconnecting HMMs and approximating an augmented problem of (11) yielding

$$\Gamma_N^* := \underset{all\ \Gamma_N}{\operatorname{argmax}} \left\{ \max_{all\ \mathcal{W}_N} P(\mathcal{O}_N, \mathcal{W}_N|\Gamma_N) P(\Gamma_N) \right\}, \quad (22)$$

where $\mathcal{O}_N$ and $\mathcal{W}_N$ are the total observation respectively state sequence over the connected models. The approximation of $P(\mathcal{O}_N|\Gamma_N)$ is necessary in practice to solve this complex problem efficiently using the Viterbi-algorithm (cf. 12-15). The prior $P(\Gamma_N) = P(\gamma^1)P(\gamma^2|\gamma^1)\ldots P(\gamma^N|\gamma^1, \ldots, \gamma^{N-1})$ (also known as *language model* in speech recognition) may be used to statistically model contextual information. E.g. a player of a *pervasive game* [9] might visit some (sub-)rooms more often (with higher probability) than other places of the whole game environment dependent on previously visited (sub-)rooms and steered by the gameplay.

3. Experiments and System Setup

In the previous sections, the hardware implementation as well as the principle of stochastic modeling inside the SYLOC-system were discussed. In the following the experiments with the SYLOC prototype and the system configuration will be explained. The experimental prototype is currently not capable of real-time operation, thus evaluation, training and testing was performed offline using a well-established toolkit for Hidden-Markov-models, called *Hidden-Markov-Model-Toolkit (HTK)* [10]. HTK has been mainly developed to build-up and test continuous speech recognition systems. However, it is general enough to be applicable to other applications demanding HMMs. The test environment was a floor of a building consisting of six rooms: bedroom, bathroom, hallway, kitchen, living room

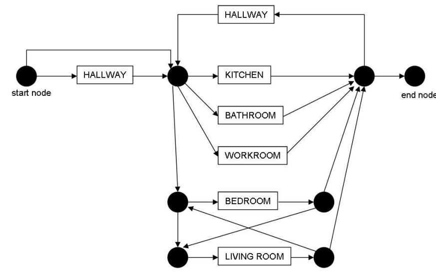

Figure 9. The shown network represents allowed paths through the environment.

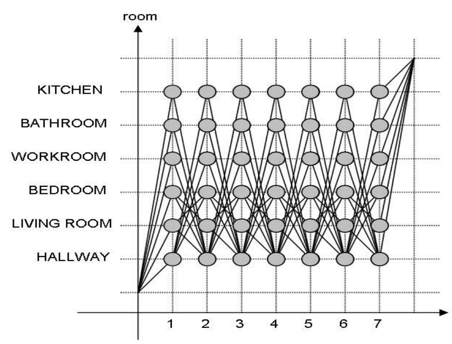

Figure 10. The Viterbi algorithm finds the most probable room sequence.

and workroom. A sketch of this environment including the most affective furnitures are shown in (Fig. 1). Obviously, not all paths through the spaces are permitted. For example, one always has to walk from room to room across the hallway while a direct jump between bedroom, bathroom, kitchen, living room and workroom is in general not possible. This knowledge may improve the performance of a recognition system and therefore it was implemented in SYLOC by a so called *task grammar* (Fig. 9). The task grammar defines a lattice of allowed room transitions and represents the global structure of the environment. The resulting lattice (Fig. 10), consisting of nodes as well as arcs between them, is used by the Viterbi algorithm to find the most probable room sequence Γ_N^* (cf. 22). The nodes may be though of as the doors of the rooms (actually, they are the start nodes respectively end nodes of the associated HMMs) whereas the arcs represent the acoustic characteristic $P(\mathcal{O}|\gamma^q)$, $(1 \leq q \leq 6)$ of a room. Note, no context information were used, thus all $P(\Gamma_N)$ were assumed equal. In the shown network without context probabilities, the probability of a path is simply the product of the observation probabilities along that path. For the acoustic modeling $\boldsymbol{b}^q$, mixtures of M Gaussians with diagonal covariance

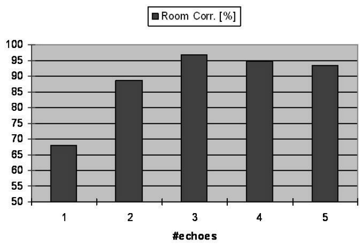

Figure 11. The diagram shows the evaluated (RC [%]) for different number of echoes d.

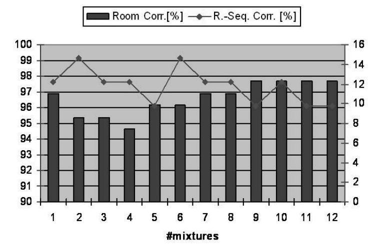

Figure 12. The diagram shows the evaluated (RC [%], left axis) and (RSC [%], right axis) involving $d = 3$ echoes and different number of mixtures.

matrices $\mathbf{\Sigma}_m^k$, $(1 \leq k \leq K^q)$, $(1 \leq m \leq M)$ has been assumed in any of the K^q states of the room-HMM γ^q. The choice of K^q was dependent on the length of the room (approx. 0.7m a state) plus a start and end state representing the doors interconnecting the HMMs (Fig. 7). The task grammar was also employed to randomly generate room sequences for the acquisition of evaluation, training and test data by randomly traversing the resulting lattice and outputting the associated label of each room start node encountered. A total of 250 observation sequences were acquired using SYLOC and composed in three disjoint sets: 50 sequences for evaluation (130 room instances), 150 sequences for training (633 room instances) and further 50 sequences for testing (154 room instances).

3.1. Training of Room-HMMs using HTK

For evaluation and testing of the SYLOC-system a HMM for every room of the test environment had to be trained using the training data. The training procedure were performed in three phases using the tools of HTK. First, the MEMU measured so called *bootstrap data* by separately traversing every room (cf. Fig. 1) in each case 30 times beginning and ending at the doors (*room boundary*). Then, the room-HMMs γ^q were initialized with a *flat start* where the bootstrap data related to the model was used to estimate its global means and variances of the diagonal covariance matrices. The initialization of the transition probabilities is uncritical and thus they were set to any value satisfying the positivity and summation constraints. From this flat start, sophisticated *start models* were reestimated using the Baum-Welch algorithm and the bootstrap data again. In the final training steps, when only training data of room sequences without any boundary information was available, the room sequence transcriptions generated from the task grammar together with the training data were processed in turn to construct a composite HMM which spans the whole sequence. The Baum-Welch formulas were then applied repeatedly until convergence had been achieved.

3.2. Evaluation of System Parameters

The evaluation data was used to experiment with adjusting the affecting parameters of the SYLOC-system independent of the test data. With increasing number of echoes d the MEMU receives, the number of free parameters to be estimated increases, too. Thus, a huge amount of data would be needed for a robust estimate of the room-HMMs. Usually, this makes a system setup very extensive perhaps even impossible for a new environment the SYLOC-system shall be employed. Obviously, only a small number of effective depth differences in the direction of sound can be expected in small (sub-)rooms with moderate number of static objects inside. Therefore, when there are many echoes they are likely to include redundant or even incorrect depth information, because of multi-echo interferences (Fig. 4). To find out an optimal d for the test environment, HMMs were trained (cf. Sec. 3.1) using the training data with a fixed number of mixtures ($M = 1$) but with varying number of echoes (Fig. 11). The optimum is found for $d = 3$ with 96.92% room recognition rate (*RC*), as expected the performance was worse when using more echoes. Setting $d = 3$, the next evaluation experiment examined the influence of the number of mixtures M in each of the observation distributions b_k (Fig. 12). As best setting, $M = 10$ has been chosen resulting in 97.69% (RC) as well as 12.00% room sequence recognition rate (*RSC*). After tuning the insertion penalty in order to balance insertion and deletion errors, the evaluation performance finally reached 90.00% (RC) and 63.41% (RSC).

3.3. Using Contextual Room Models

As shown in (Fig. 7), the hallway was modeled using one HMM only, although it can be passed through many different routes. Obviously, using the same HMM for very different paths results in bad acoustic modeling and hence many insertion and deletion errors may occur. Therefore, contex-

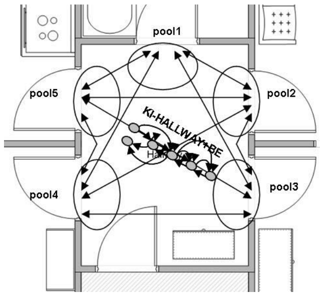

Figure 13. Possible paths modeled by room-dependent HMMs. The ellipses indicate clustered states.

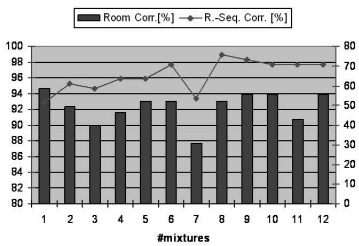

Figure 14. The diagram shows the evaluated (RC [%], left axis) and (RSC [%], right axis) using the tied tri-room models ($d = 3$).

tual HMMs were trained to account for the dependency on the predecessor and successor room when walking across the hallway (Fig. 13). Because of the dependency of the predecessor and successor such models are called *tri-room models* as opposed to *mono-room models* used so far.

3.4. Tied Tri-Room Model Training and Evaluation

In principle, the training of tri-room models for the hallway is similar to the mono-model training, but in the presence of very limited data one has to establish a more elaborate training procedure: First, 10 tri-room models *-HALLWAY+* named according to the successor and predecessor room [1] were cloned from the HALLWAY mono-room model ($M = 1$) (cf. Sec. 3.1). No distinction was made between forward and backward paths. Models with similar path structure were forced to share the same transition matrices to reduce complexity. The resulting tri-room models [2] were then reestimated with mono-room labels substituted by tri-room labels. To reduce the model complexity further, eight pools of acoustic similar positions (states) were created and the associated states of a pool were shared. After reestimation, the remaining HALLWAY labels with unknown predecessor were substituted by bi-labels HALLWAY+* and in turn a Viterbi automated realignment was performed to substitute these by the most probable tri-room models. Evaluation results using all reestimated tied tri-room models are shown in (Fig. 14). The models ($M = 8$) performing 75.61% (RSC) as well as 93.06% (RC) were chosen for the final test.

[1] The symbol * means any of the abbreviations {BE,BA,KI,LI,WO} for {BEDROOM,BATHROOM,KITCHEN,LIVING ROOM,WORKROOM}

[2] For the sake of brevity, all models are called tri-room models, although the hallway is the only contextual room.

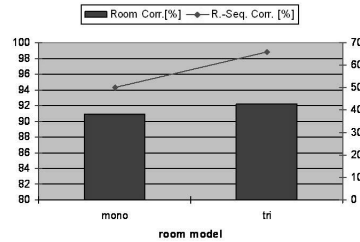

Figure 15. The figure shows the (RC [%], left axis) and (RSC [%], right axis) obtained on the test data.

3.5. Testing

After configuring the system parameters to the evaluated values, the final step was an independent test of the SYLOC-system using the test data which has been retained so far. The trained mono-room models (Sec. 3.1) and tri-room-models (Sec. 3.4) were separately employed to recognize the test data and to measure the performance by means of different figures of merit. In (Fig. 15) the (RC) and (RSC) on the test data are shown. The use of tri-room models improved the system performance by a relative (RC) of 1.43% whereas (RSC) increased by 32.00% relating to the mono-room models. The absolute rates are 92.21% (RC) and 66.00% (RSC) for tri-room models and 90.91% (RC) and 50.00% (RSC) for mono-room models. The improvement of the (RSC) is due to the decrease of the insertion and deletion errors as a result of a better acoustic modeling with tri-room models for the hallway (Fig. 16). In (Fig. 17) a confusion matrix is shown to give a detail description of the performance of the SYLOC-system in recognizing the different rooms using the tri-room models. The vertical represents the correct category, whereas the horizontal represents the classification of the SYLOC-system. The performance is very good, with only 2 incorrect classifications of WORKROOM as LIVING ROOM. The room sequence recognition rate is mainly affected by the deletions and insertion errors of the hallway.

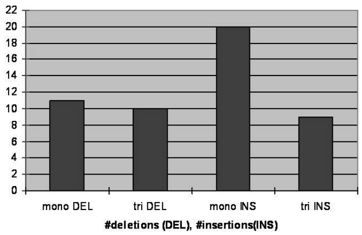

Figure 16. The figure shows the deletion and insertion errors.

	HA	KI	WO	BA	BE	LI	DEL
HA	52	0	0	0	0	0	10
KI	0	15	0	0	0	0	0
WO	0	0	10	0	0	2	0
BA	0	0	0	12	0	0	0
BE	0	0	0	0	25	0	0
LI	0	0	0	0	0	28	0
INS	9	0	0	0	0	0	

Figure 17. The confusion matrix. Additionally, deletion and insertion errors are given.

4. Conclusion

This paper presented a symbolic localization system (SYLOC) based on topological features extracted from ultrasonic multi-echo range measurements in four coplanar and orthogonal directions in space. SYLOC used HMMs to estimate positions by means of a classification of continuous acquired features to (sub-)room symbol sequences. Promising results of 92.21% room recognition rate and 66.00% room sequence rate are presented and focus further research in developing a next SYLOC prototype. The additional effort using tied tri-room models is marginal, as cloning of tri-room models from mono-room-models is automated and the rules for pooling of states must be defined only once. Although the idea of using ultrasonic ranging for navigation is not new (especially in robotics), the novel concept for personal tracking presented in this paper of using continous multi-echo measurements in connection with HMMs has many advantages compared to ultrasonic based systems reported in literature [1, 2, 4, 5, 6]:

- SYLOC is self-contained. No preliminary installed equipment is necessary, because ranging is initiated from the moving user in opposite to measurements performed from the walls or the ceiling by fixed sensors.
- SYLOC is mobile for personal tracking.
- SYLOC is a low cost solution, only commercial inexpensive hardware devices are used. For the algorithmic implementation existing (free available) tools were employed for training HMMs as well as Viterbi decoding.
- SYLOC is easy adaptable to changing environments using well-known adaptation techniques, like e.g. MLLR or MAP approach which are also well-known from continuous speech recognition.

References

[1] A. Ward, A. Jones and A. Hopper, *A New Location Technique for the Active Office.* IEEE Personal Communications 4, 5:43-47, October 1997

[2] N.B. Priyantha, A. Chakraborty and H. Balakrishnan, *The Cricket Location-Support System.* 6th ACM Int. Conf. on Mobile Computing and Networking, Boston, August 2000

[3] P. Bahl and V.N. Padmanabhan, *Radar: An In-building RF-based User Location and Tracking System.* Proceedings of the IEEE Infocom, Tel Aviv, Israel, 2:775-784, 2000

[4] H.W. When and P.R. Belanger. *Ultrasound-based robot position estimation.* IEEE Transaction on Robotics and Automation, 13(5):682-692, 1997

[5] A.R. Jiménez and F. Seco, *Ultrasonic Localization Methods for Accurate Positioning*, New Acoustics : selected topics II, Colección Biblioteca de Ciencias del CSIC, 167-186, 2006

[6] O. Aycard, J.-F. Mari and R. Washington, *Learning to Automatically Detect Features for Mobile Robots Using Second-Order Hidden-Markov-Models.* Int. Journal of Advanced Robotic Systems, 1, 4:231-244, 2004

[7] Devantech Ltd (Robot Electronics), Unit 2B, Gilray Road, Norfolk, England

[8] L.R. Rabiner, *A Tutorial on Hidden-Markov-Models and selected Applications in Speech Recognition* Proceedings of the IEEE, 77, 2:257-286, February 1989

[9] W. Broll, J. Ohlenburg, I. Lindt, I. Herbst and A.-K. Braun, *Meeting Technology Challenges of Pervasive Augmented Reality Games.* ACM NetGames 2006 - 5th Workshop on Network & System Support for Games, 2006, CD-ROM Proc. also published in ACM Digital Library

[10] S. Young and P. Woodland, *HTK - Hidden Markov Model Toolkit*, Cambridge University Engineering Department (CUED), http://htk.eng.cam.ac.uk/