Using the Support Vector Regression Approach to Model Human Performance

Luzheng Bi, *Member, IEEE*, Omer Tsimhoni, *Member, IEEE*, and Yili Liu, *Member, IEEE*

***Abstract*—Empirical data modeling can be used to model human performance and explore the relationships between diverse sets of variables. A major challenge of empirical data modeling is how to generalize or extrapolate the findings with a limited amount of observed data to a broader context. In this paper, we introduce an approach from machine learning, known as support vector regression (SVR), which can help address this challenge. To demonstrate the method and the value of modeling human performance with SVR, we apply SVR to a real-world human factors problem of night vision system design for passenger vehicles by modeling the probability of pedestrian detection as a function of image metrics. The results indicate that the SVR-based model of pedestrian detection shows good performance. Some suggestions on modeling human performance by using SVR are discussed.**

***Index Terms*—Human factors, human performance data analysis and modeling, night vision systems, pedestrian detection, support vector regression (SVR).**

I. Introduction

HUMAN performance modeling has several potential benefits for the study of human–machine interaction and system design. Such modeling can help predict human performance, reduce cost, and guide system design. Much attention has been devoted to this area of research [1], and a wide range of models has been developed, from performance level (such as models of reaction time, movement time, and detection probability) to processes underlying human performance (such as detailed representations of cognitive processes) [2]–[6]. Human performance models include both theoretical modeling and empirical data modeling (often called empirical modeling). Usually, theoretical modeling starts with and tests certain theories regarding human performance or human–machine interaction, which are then revised and refined on the basis of empirical data. Empirical modeling, on the other hand, is usually more data driven from the beginning and proceeds by finding the best mathematical method to develop a quantitative function between human performance and empirical data and variables.

Manuscript received August 27, 2008; revised December 16, 2009; accepted June 16, 2010. Date of publication October 24, 2010; date of current version April 15, 2011. This paper was recommended by Associate Editor A. Bargiela.

L. Bi is with the School of Mechanical Engineering, Beijing Institute of Technology, Beijing 100081, China (e-mail: bhxblz@bit.edu.cn).

O. Tsimhoni is with the General Motors Advanced Technical Center—Israel, Herzliya 46725, Israel (e-mail: omer.tsimhoni@gm.com).

Y. Liu is with the Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, MI 48109-2117 USA (e-mail: yililiu@umich.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSMCA.2010.2078501

Theoretical modeling can help understand the detailed processes of human performance and human–machine interaction, as illustrated in the development of cognitive architectures such as adaptive control of thought—rational [6], [7] and queuing networks [2]–[5] and some earlier models based on control theory or information theory [8]–[10]. Comprehensive reviews of the broad field of human performance modeling can be seen in, e.g., [11]. In many situations, however, corresponding theories may not exist, and thus, empirical modeling is needed to model human performance and solve the related specific practical problems. The approach may also offer insights into the domain of inquiry and may serve as a basis for future theory development.

A major challenge of empirical modeling is how to generalize or extrapolate the findings with a limited amount of observed data to a broader context. This challenge arises because it is often impossible or impractical to collect data on all the possible scenarios of a human–machine system application or over the full range of possible values of a crucial design or operational variable. Researchers and practitioners often need to extrapolate or generalize their findings beyond the scope of their empirical study.

In human factors, the most widely adopted method for investigating and modeling the relationships among empirical variables is the family of linear and polynomial least squares regression methods. These methods are widely used to analyze empirical data and to model human performance as a function of various variables by establishing the best fit of a model to the observed data [12]–[16]. However, these methods have limited ability to approximate nonlinear relationships and thus cannot provide a good fit of the observed nonlinear data. Furthermore, the goodness of fit in one data set does not necessarily mean the goodness of fit in other data sets, particularly when these data fall outside the range of the observed original data. Therefore, mathematical techniques need to be identified or established as more valid support for extrapolation and generalization purposes in human factors research.

Interestingly, the field of machine learning faces a similar challenge in its research on the relationship between training data and other nontraining data. In machine learning, in addition to conventional regression methods, artificial neural network (ANN) is a widely used nonlinear modeling method for regression analysis and classification [17]–[24]. This method is based on biological neural networks. The main advantage of ANN as a regression analysis method is its ability to approximate a continuous function by minimizing the error in fitting a limited amount of observed data [25]. In machine learning,

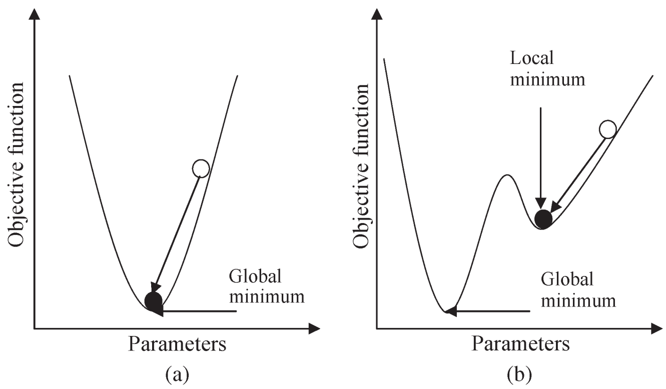

Fig. 1. Objective functions of the parameters of (a) SVR and (b) conventional nonlinear least squares regression and ANN.

this error is called training error, and the set of observed data is called training data.

How do we know which method can be used to develop the best model to approximate the real model? Since it is impossible for us to know the real model, we cannot evaluate these methods by comparing the developed models with the real model. A main criterion of evaluating these developed models is their generalization performance, which refers to a model's accuracy in predicting output values based on input variables of other data (future data) rather than training data. The goodness of fit in modeling the training data does not guarantee the goodness of fit in other data due to the presence of noise, which is often called overfitting [26]. In machine learning, the error of models in predicting future data is called generalization error or testing error.

Recently, in machine learning, a new regression method, known as support vector (SV) regression (SVR), has been developed and applied to help improve generalization performance [27]–[35]. SVR is an application of SV machines (SVMs) in regression analysis. SVM is a group of supervised learning methods that can be applied to classification or regression, and uses statistical learning theory as a general mathematical framework with a limited amount of data [36], [37]. Unlike conventional regression methods and ANN, SVR develops models by minimizing the upper bound of the generalization error rather than minimizing the training error [38]. This gives SVR a greater ability to generalize (we describe an example to illustrate the overfitting issue and this ability of SVR in Section II-C). Improving generalization performance and minimizing the error in fitting training data are both optimization processes, whose success depends on the characteristics of the employed optimization functions. The objective function of the parameters of SVR-based models is a quadratic function as shown in Fig. 1(a), which only has one global minimum. Thus, most optimization parameters of SVR-based models can be searched accurately. In other words, SVR does not suffer from the problem of "being trapped" at local minima. Compared to SVR, the objective functions of the parameters of conventional nonlinear regression and ANN are likely to have a shape as shown in Fig. 1(b) and can possibly trap the search process into a local minimum. In other words, the search process can possibly never find the global minimum. Table I reviews some characteristics of the methods discussed earlier.

Generalization performance is an important factor in evaluating the different regression methods. The basic idea to estimate the generalization performance of one method is that testing data and training data should be different. In practice, K-fold cross validation is widely used [39]–[42]. Generally, R^2 (coefficient of determination) and root-mean-square (rms) error (rmse) are selected as performance indicators. The greater is the R^2 value, the greater is the proportion of the variation in the dependent variables that is explained by the models. Similarly, the smaller is the value of rmse, the smaller is the prediction error of the models, representing better performance.

In the following, we first introduce the basic principle of SVR, as well as the SVR-based modeling approach. Then, to demonstrate how to apply SVR-based modeling in human factors research, we use SVR to model the probability of pedestrian detection with night vision systems. Consequently, we compare the generalization performance of the SVR-based model with a model based on Stevens' law, along with some models developed with ANN, polynomial regression, and linear regression. Finally, we discuss the importance of the SVR-based modeling approach as an option for human performance modeling and offer some suggestions on how to use SVR to model human performance.

II. SVR-Based Modeling

A. Basic Principle of SVR

Suppose that we are given a training data set

$$\{x_i, y_i\}, \qquad i = 1, \ldots, l; \quad x_i \in R^d; \quad y_i \in R^d \tag{1}$$

where x_i is a vector of input variables and y_i represents the corresponding scalar output (target) value. The goal of SVR-based modeling is to develop a function $f(x)$ that can accurately predict the output values based on their input variables on future data. SVR performs the goal through finding a function called ε-SVR function that has at most ε deviation from the actual obtained output values for all the training data and that is as flat as possible (for controlling the function capacity) [43].

In the case of linear functions, $f(x)$ can be

$$f(x) = \omega \bullet x + b \tag{2}$$

where $\bullet$ denotes the dot product, ω is the weight vector, and b refers to the constant (bias). Flatness in the case of (2) can be achieved by minimizing the Euclidean norm of the weight vector ω, i.e., $\|\omega\|$. Formally, this problem can be expressed as a convex optimization problem

$$\begin{aligned} &\text{Minimize} \quad \frac{1}{2}\|\omega\|^2 \\ &\text{Subject to} \quad \begin{cases} y_i - f(x_i) \le \varepsilon \\ f(x_i) - y_i \le \varepsilon. \end{cases} \end{aligned} \tag{3}$$

Sometimes, however, such a function that fits all the observed data with ε precision does not exist. Correspondingly, the slack

TABLE I
COMPARISON OF EMPIRICAL MODELING METHODS

	SVR	ANN	Polynomial regression	Linear regression
Ability to model nonlinear relation	Strong	Strong	Moderate	No
Robust to local minima	Yes	No	No	Yes
Optimization criterion	Upper bound of generalization error	Training error	Training error	Training error
Required level of expertise	Expert	Expert	Moderate	Easy

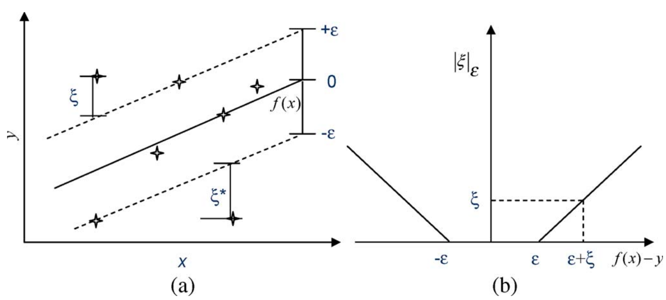

Fig. 2. (a) Loss setting for linear SVR and (b) ε-insensitive loss function.

variables ξ_i and ξ_i^* are introduced to address the infeasible constraints of the optimization problem (3). SVR not only minimizes the training error by minimizing the sum of ξ_i and ξ_i^* but also minimizes $\|\omega\|$ in order to increase the flatness of the function. This optimization problem can be expressed as [36]

$$\text{Minimize} \quad \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{l}(\xi_i + \xi_i^*)$$
$$\text{Subject to} \quad \begin{cases} y_i - f(x_i) \le \varepsilon + \xi_i \\ f(x_i) - y_i \le \varepsilon + \xi_i^* \\ \xi_i^*, & \xi_i \ge 0; \quad C > 0 \end{cases} \tag{4}$$

where C is a constant determining the tradeoff between the flatness of the function and the training error. This is associated with handling a so-called ε-insensitive loss function

$$|\xi|_\varepsilon = \begin{cases} 0, & |f(x) - y| \le \varepsilon \\ |f(x) - y| - \varepsilon, & \text{otherwise.} \end{cases} \tag{5}$$

Fig. 2(a) shows the loss setting graphically, and Fig. 2(b) shows the ε-insensitive loss function. One main reason for using the ε-insensitive loss function is that it can produce a sparse set of SVs, which makes practical computation feasible. In contrast, other loss functions (such as Huber, Gaussian, and Laplace loss functions) do not produce a sparse set of SVs, making the computation difficult, if not infeasible. A more detailed analysis and discussion of these different loss functions can be found in [44].

The convex optimization problem (4) can be solved more easily in its dual form in most cases. After (4) is transformed into a dual objective function, it takes the following form [44]:

$$\text{Maximize} \quad -0.5\sum_{i,j=1}^{l}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)(x_i \bullet x_j) + \sum_{i=1}^{l} y_i(\alpha_i - \alpha_i^*) - \varepsilon\sum_{i=1}^{l}(\alpha_i + \alpha_i^*)$$
$$\text{Subject to} \quad \sum_{i=1}^{l}(\alpha_i - \alpha_i^*) = 0 \qquad 0 \le \alpha_i, \qquad \alpha_i^* \le C \tag{6}$$

where α_i and α_i^* are Lagrange multipliers for the ith training example and are gained by solving (6). Only some of the coefficients $(\alpha_i - \alpha_i^*)$ are nonzero, and the corresponding input vectors x_i are called SVs. These SVs x_i and the corresponding nonzero Lagrange multipliers α_i and α_i^* give the values of weight vector ω, which can be expressed as

$$\omega = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)x_i. \tag{7}$$

By combining (2) with (7), we derive

$$f(x) = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)(x \bullet x_i) + b. \tag{8}$$

Also, the bias parameter b can be computed by applying Karush–Kuhn–Tucker conditions [37]. Equation (8) is the linear model developed with this method.

However, in most practical cases, the relationship between input variables and output values is not linear. The SVR-based nonlinear model can be developed by simply mapping input variables into a high-dimensional feature space F (i.e., by a map $\Phi : R^d \to F$) [36]. The nonlinear function is formed as follows:

$$f(x) = \omega \bullet \phi(x) + b. \tag{9}$$

With the same process of solving the linear function, we derive

$$f(x) = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)(\phi(x) \bullet \phi(x_i)) + b. \tag{10}$$

However, as the input dimensions increase, the dimensions in the feature space further increase by many folds, and thus, the mapping process becomes a computationally infeasible problem. This problem can be dealt with by defining appropriate kernel functions in place of the dot product of the input vectors in the high-dimensional feature space [36]. The kernel function is expressed as

$$K(x, x_i) = \phi(x) \bullet \phi(x_i). \tag{11}$$

Combining (10) with (11), we derive the regression function that takes the following general form:

$$f(x) = \sum_{i=1}^{l} (\alpha_i - \alpha_i^*) K(x, x_i) + b. \tag{12}$$

Currently, several kinds of kernel functions have been defined and used, such as linear functions, polynomial functions, and radial basis functions (RBFs). Among these, the RBF, as defined in the following, is most commonly used:

$$K(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right) \tag{13}$$

where $1/2\sigma^2$ represents the width of the RBF.

More details about SVR can be found in a recent tutorial on SVR [44].

B. Implementation Method

In practice, one does not need to solve mathematically the complex optimization problem in order to use SVR but can instead use software packages such as the LIBSVM software library developed by Chang and Lin [45]. The LIBSVM (library for support vector machines) package is used to develop SVR-based models. It uses a fast and efficient method known as sequential minimal optimization for solving large quadratic programming problems. The detailed procedure of developing an SVR-based model by using the LIBSVM software can be found in [45].

After the model is built, it can be used to predict the target value based on input variables of new data. This can be implemented with the command "svmpredict" in the LIBSVM package [45].

C. Example

To illustrate that the goodness of fit in the training data does not mean the goodness of fit in other data and that SVR tends to show a better ability to generalize, we estimated functions from a set of artificial data shown in Fig. 3, by using linear regression, polynomial regression, ANN, and SVR, respectively. We then employed a threefold cross validation to evaluate the generalization performance of these different methods.

First, the collected data set was randomly divided into three partitions (subsets). Second, three models for each of the four modeling methods were built. Each model used two subsets of the data as a training set by leaving out one subset each time. Finally, the average of the errors associated with the left-out subsets was used as an estimate of the generalization error, whereas the average of the errors associated with the training sets was used as an estimate of the training error.

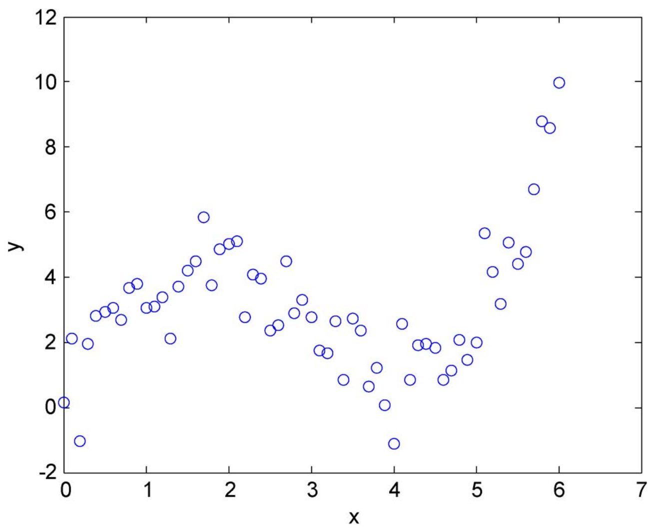

Fig. 3. Set of artificial data for evaluating the generalization performance of several methods.

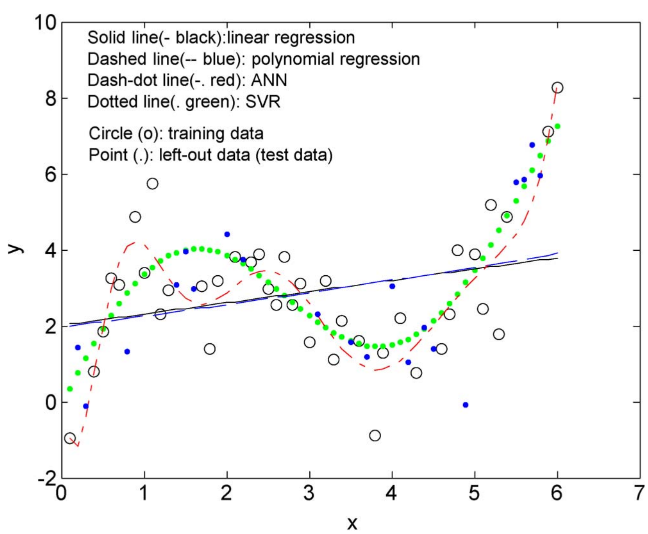

Fig. 4. Functions developed by four regression methods on one training data set.

RBF was used as the kernel function of SVR. By evaluating the SVR-based model for a wide range of the width of the RBF kernel $(1/2\sigma^2)$, cost coefficient (C), and loss function parameter (ε), the final parameters were determined: $1/2\sigma^2 = 0.1, C = 100$, and $\varepsilon = 1$. The three models based on SVR were then each obtained with one of the three training sets according to these final parameters by using "svmtrain" in the LIBSVM package. The hide function of ANN was taken as RBF for the comparison with SVR-based models.

The three training data sets were used to build three groups of models. Each group contains four models, built with the four regression methods. Fig. 4 shows one group of the models.

Fig. 5 shows the generalization performance $(+)$ and the training performance $(\circ)$ corresponding to each method with

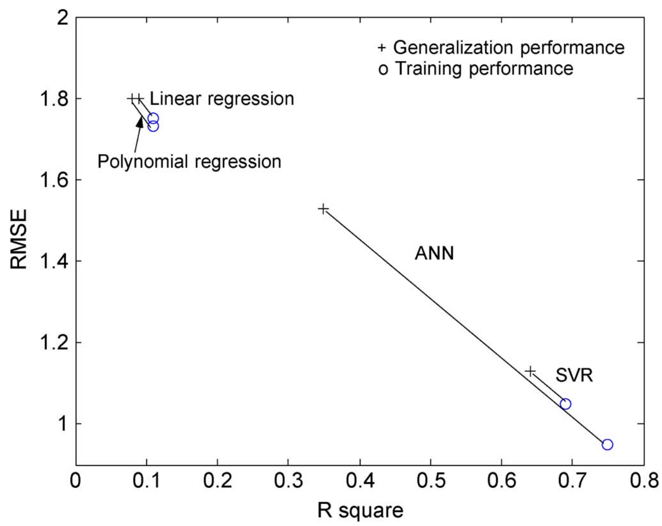

Fig. 5. (+) Generalization performance and (○) training performance of each method.

the two performance indicators rmse and R^2 (higher values of R^2 and lower values of rmse represent better performance).

As shown in Fig. 5, ANN has the best training performance. It does not, however, have the best generalization performance. This is because ANN has a strong ability to approximate continuous functions by minimizing the training error, and thus, ANN overfits the training data. Although SVR does not show the best training performance, it has the best generalization performance. This is because SVR can make a tradeoff between the training error and its ability to approximate continuous functions so as to minimize the upper bound of the generalization error rather than minimizing the training error, and thereby avoid overfitting. Linear and polynomial regressions show worse training performance than SVR and ANN because they do not have enough ability to fit the training data. In addition, linear and polynomial regressions do not have good generalization performance in this case. One reason for this is that, according to [26], the generalization error of the functions is equal to or less than the sum of the training error and a nonnegative term, and linear and polynomial regressions show worse training performance in this case. Another possible reason is that the objective function of the parameters of polynomial regression did find the local minimum rather than the global minimum, as discussed in Section I.

III. Modeling the Probability of Pedestrian Detection With Night Vision Systems

In this section, we apply the SVR method to a real-world human factors challenge in the context of night vision system design and evaluation for passenger vehicles. The primary objective of night vision systems in passenger vehicles is to help drivers detect pedestrians in the dark [46]. In night vision system design and evaluation, the main objective of developing models is to help researchers estimate pedestrian detection performance without spending a significant amount of time designing experimental materials, recruiting subjects, and conducting experiments. Because the physical characteristics of images generated by night vision systems can be measured by image-based metrics, which greatly impact detection performance, pedestrian detection performance can be modeled as a function of image metrics with these night vision systems.

Several studies have investigated modeling the relation between pedestrian detection distance and the metrics of images generated by night vision systems. Tsimhoni *et al.* [47] found a high correlation between mean detection distance and subjective rating of image clutter. Lim *et al.* [48] modeled the mean detection distance based on clutter metrics of images and found that, when the clutter metric is lower, the detection distance to pedestrians is longer. In addition, several models have been developed to predict the detection probability of military targets in infrared images based on image metrics, and these models show good performance [49]–[53]. However, these models were developed for images in a military context and were not focused on pedestrian detection. To better understand the probability of pedestrian detection, Bi *et al.* [54] have recently developed a model of the probability of pedestrian detection based on image metrics by using Stevens' law as the theoretical framework and guidance (this approach, therefore, falls into the theoretical modeling category as described earlier in this paper). In the following, we model the probability of pedestrian detection as a function of image metrics using SVR and compare the modeling results with those of the model based on Stevens' law and the models developed by linear regression, polynomial regression, and ANN, respectively.

A. Image Metrics

As we mentioned earlier, the physical characteristics of images generated by night vision systems can be measured by image-based metrics. Generally, these image metrics include clutter metrics, contrast metrics, and target size metrics.

1) Clutter Metrics: We selected one commonly used clutter metric in our modeling work. The rms clutter metric, proposed by Schmieder and Weathersby [49], is a classic clutter metric of the spatial intensity of an image. The image considered is first divided into N adjacent square blocks. The size of each block is defined as twice the size of the largest target dimension. Then, the individual intensity variances are computed for each block. Finally, the rms clutter metric is calculated by averaging the variances of all the blocks. The rms clutter metric is expressed as follows:

$$rms = \left(\frac{1}{N} \sum_i \sigma_i^2 \right)^{1/2} \quad (14)$$

where σ is the individual intensity standard deviation within each block.

2) Contrast Metrics: Contrast metrics measure the intensity difference between a target and its local background. We used one commonly used and easily calculated contrast metric, the Doyle metric [50], which measures the differences in the means and variances between the target and local background intensities. It is defined as

$$Doyle = \left[(\mu_T - \mu_B)^2 + (\sigma_T - \sigma_B)^2 \right]^{1/2} \quad (15)$$

Fig. 6. One frame from an FIR video clip.

Fig. 7. One frame from an NIR video clip.

where μ_T and μ_B represent the mean intensities of the target and of the background, respectively, and σ_T and σ_B represent the standard deviations of the target and of the local background, respectively.

3) Target Size Metrics: A target size metric may be defined as the square root of the number of pixels on a target (RPOT). It is expressed as

$$RPOT = \sqrt{POT} \tag{16}$$

where POT is the number of pixels on a target.

B. Collection of the Data Set

To compare the SVR-based model with the other models, we used the same visual images and data on pedestrian detection performance from Tsimhoni *et al.* [47], who examined the relation between pedestrian detection performance and two types of night vision systems—far-infrared (FIR) systems and near-infrared (NIR) systems. Both FIR and NIR systems were used in this study for the purpose of collecting images with a wide range of image metric values as a data source to support modeling of detection performance based on image metrics.

Of the 113 video frames collected, 62 were from FIR system video clips, and 51 were from NIR system video clips. The video clips came from recorded images of naturalistic driving (more details can be found in [47]). Figs. 6 and 7 show one frame from an FIR system video clip and one frame from an NIR system video clip, respectively. By implementing (14)–(16), the values of the metrics over all the frames were obtained.

A total of 16 licensed drivers—eight younger (ages 21 to 30, mean of 24, four women and four men) and eight older (ages 64 to 79, mean of 71, four women and four men)—participated in a target task experiment. The detection probabilities associated with all the frames were collected. In this study, detection probability is defined as the number of subjects who correctly indicate the presence of the pedestrian on the frame over the number of all subjects. In other words, it is defined as the percentage of subjects who correctly detected the presence of a pedestrian.

The detailed procedure of the experiment is reported in [47].

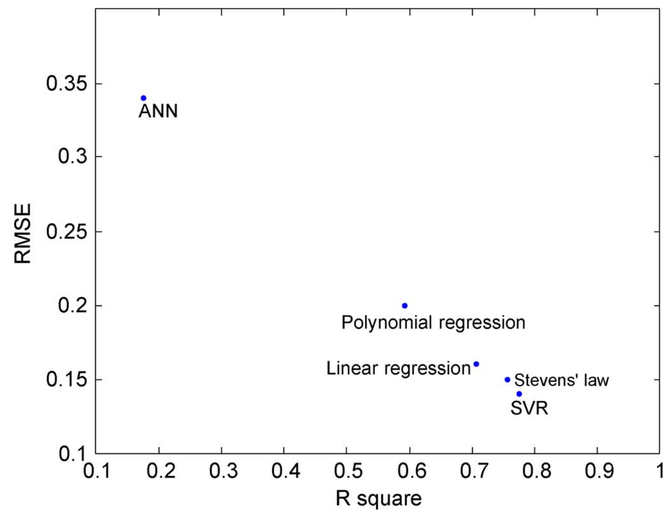

Fig. 8. Generalization performance of the five methods.

C. Modeling the Probability of Pedestrian Detection and Results

After the values of the image metrics and detection probabilities were collected, we took the rms, RPOT, and Doyle metrics as input variables to model the probability of pedestrian detection by using SVR and other methods.

We first used a fourfold cross validation to estimate the generalization performance of these different methods. The collected data were randomly divided into four partitions (the sample amounts for each partition were 28, 28, 28, and 29). The kernel function of SVR was selected as the RBF kernel. By evaluating the SVR-based model for the wide range of parameters [including the width of the RBF kernel $(1/2\sigma^2)$, the cost coefficient (C), and the loss function parameter (ε)], the final parameters were determined: $1/2\sigma^2 = 1$, $C = 10$, and $\varepsilon = 0.15$. According to the values of these parameters, the final SVR-based model can be built with the collected data by using the command "svmtrain" in LIBSVM. Furthermore, we applied the fourfold cross-validation method to estimate the performance of the models developed by Stevens' law, linear regression, polynomial regression, and RBF ANN.

The results of the two performance indicators of all the models as a function of the rms, RPOT, and Doyle metrics are shown in Fig. 8. The performance of the SVR-based model is almost identical to, even slightly better than, that of the model developed on the theoretical basis of Stevens' law. The linear regression model shows relatively good performance in this case, very possibly because the relationship between detection probability and the three metrics is close to linear, in contrast to the strongly nonlinear relationship between variables shown in the illustrative data set in Fig. 4.

IV. Conclusion

In this paper, we have described an empirical modeling method called SVR, which can help improve a model's generalization performance. We used SVR to model the probability of pedestrian detection with night vision systems as a function of image-based metrics and compared the SVR-based model with the model developed on the theoretical basis of Stevens' law, along with other models developed with linear regression, polynomial regression, and ANN. The results indicate that the SVR-based model shows the same level of performance as the model developed on the theoretical basis of Stevens' law and somewhat better performance than the three empirical modeling methods.

This is because SVR can make a tradeoff between minimizing the training error and its ability to approximate continuous functions so as to minimize the upper bound of the generalization error. This gives SVR a greater ability to generalize.

Although SVR-based modeling and other empirical modeling methods do not typically specify the detailed processes of human–machine interaction, the generalized predictions of human performance achieved with SVR are helpful and valuable in supporting system design.

In practice, SVR-based modeling may be perceived for some modelers as too difficult to use. However, several software tools exist (such as the LIBSVM package) that can be used quite easily for this purpose. Thus, we suggest that, for better generalization performance in human factors modeling and data analysis, SVR-based modeling should be considered an important option, particularly when theoretical performance modeling cannot be adopted due to the lack of the corresponding theories that have been satisfactorily developed. A feasible method is to try linear regression and polynomial regression first; if their generalization performance is not satisfactory, SVR-based modeling can then be utilized.

References

[1] L. Allender, "Modeling human performance: Impacting system design, performance, and cost," in *Proc. Mil., Gov. Aerosp. Simul. Symp., Adv. Simul. Technol. Conf.*, Washington, DC, 2000, pp. 139–144.

[2] Y. Liu, "Queuing network modeling of elementary mental processes," *Psychol. Rev.*, vol. 103, no. 1, pp. 116–136, Jan. 1996.

[3] C. Wu, Y. Liu, and C. Walsh, "Queueing network modeling of a real-time psychophysiological index of mental workload—P300 in evoked brain potential (ERP)," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 38, no. 5, pp. 1068–1084, Sep. 2008.

[4] Y. Liu, R. Feyen, and O. Tsimhoni, "Queueing network-model human processor (QN-MHP): A computational architecture for multitask performance," *ACM Trans. Hum. Comput. Interact.*, vol. 13, no. 1, pp. 37–70, Mar. 2006.

[5] Y. Liu, "QN-ACES: Integrating queuing network and ACT-R, CAPS, EPIC, and soar architectures for cognitive modeling," *Int. J. Hum. Comput. Interact.*, vol. 25, no. 6, pp. 554–581, Aug. 2009.

[6] J. R. Anderson, M. Matessa, and C. Lebiere, "ACT-R: A theory of higher level cognition and its relation to visual attention," *Hum. Comput. Interact.*, vol. 12, no. 4, pp. 439–462, Dec. 1997.

[7] J. R. Anderson, D. Bothell, M. D. Byrne, S. Douglass, C. Lebiere, and Y. L. Qin, "An integrated theory of the mind," *Psychol. Rev.*, vol. 111, no. 4, pp. 1036–1060, Oct. 2004.

[8] A. Tustin, "The nature of the human operator's response in manual control and its implication for controller design," *J. Inst. Elect. Eng.*, vol. 94, pp. 190–201, 1947.

[9] R. J. Jagacinski and J. M. Flach, *Control Theory for Humans: Quantitative Approaches to Modeling Performance*. Mahwah, NJ: Lawrence Erlbaum, 2003.

[10] R. C. Conant, "Laws of information which governs systems," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-6, no. 4, pp. 240–255, Apr. 1976.

[11] K. A. Gluck and R. Pew, *Modeling Human Behavior With Integrated Cognitive Architectures: Comparison, Evaluation, and Validation*. Mahwah, NJ: Lawrence Erlbaum, 2005.

[12] J. H. Lim and Y. Liu, "Modeling the influences of cyclic top-down and bottom-up processes for reinforcement learning in eye movements," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 39, no. 4, pp. 706–714, Jun. 2009.

[13] M. Bauerly and Y. L. Liu, "Computational modeling and experimental investigation of effects of compositional elements on interface and design aesthetics," *Int. J. Hum.-Comput. Stud.*, vol. 64, no. 8, pp. 670–682, Aug. 2006.

[14] A. Borowsky, D. Shinar, and Y. Parmet, "The relation between driving experience and recognition of road signs relative to their locations," *Hum. Factors*, vol. 50, no. 2, pp. 173–182, Apr. 2008.

[15] M. A. Elliott, C. J. Baughan, and B. F. Sexton, "Errors and violations in relation to motorcyclists' crash risk," *Accid. Anal. Prev.*, vol. 39, no. 3, pp. 491–499, May 2007.

[16] S. Nakahara, T. Kawamura, M. Ichikawa, and S. Wakai, "Mathematical models assuming selective recruitment fitted to data for driver mortality and seat belt use in Japan," *Accid. Anal. Prev.*, vol. 38, no. 1, pp. 175–184, Jan. 2006.

[17] Z. Liu, R. E. Torres, N. Patel, and Q. Wang, "Further development of input-to-state stabilizing control for dynamic neural network systems," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 38, no. 6, pp. 1410–1433, Nov. 2008.

[18] E. Asua, V. Etxebarria, and A. Garcia-Arribas, "Neural network-based micropositioning control of smart shape memory alloy actuators," *Eng. Appl. Artif. Intell.*, vol. 21, no. 5, pp. 796–804, Aug. 2008.

[19] A. Cevik, M. A. Kutuk, A. Erklig, and I. H. Guzelbey, "Neural network modeling of arc spot welding," *J. Mater. Process. Technol.*, vol. 202, no. 1–3, pp. 137–144, Jun. 2008.

[20] E. Oguz, B. Keskinler, and A. Tortum, "Determination of the apparent ozonation rate constants of 1:2 metal complex dyestuffs and modeling with a neural network," *Chem. Eng. J.*, vol. 141, no. 1–3, pp. 119–129, Jul. 2008.

[21] M. Peer, M. Mahdyarfar, and T. Mohammadi, "Evaluation of a mathematical model using experimental data and artificial neural network for prediction of gas separation," *J. Natural Gas Chem.*, vol. 17, no. 2, pp. 135–141, Jun. 2008.

[22] E. Uncuoglu, M. Laman, A. Saglamer, and H. B. Kara, "Prediction of lateral effective stresses in sand using artificial neural network," *Soils Found.*, vol. 48, no. 2, pp. 141–153, Apr. 2008.

[23] R. S. T. Lee, "iJADE stock advisor: An intelligent agent based stock prediction system using hybrid RBF recurrent network," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 34, no. 3, pp. 421–428, May 2004.

[24] S. J. Wu, N. Gebraeel, M. A. Lawley, and Y. Yih, "A neural network integrated decision support system for condition-based optimal predictive maintenance policy," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 37, no. 2, pp. 226–236, Mar. 2007.

[25] S. Haykin, *Neural Networks: A Comprehensive Foundation,* 2nd. Englewood Cliffs, NJ: Prentice-Hall, 1998.

[26] O. Bousquet, S. Boucheron, and G. Lugosi, "Introduction to statistical learning theory," in *Advanced Lectures on Machine Learning*. New York: Springer-Verlag, 2004, pp. 169–207.

[27] M. Cimen, "Estimation of daily suspended sediments using support vector machines," *Hydrol. Sci. J.*, vol. 53, no. 3, pp. 656–666, Jun. 2008.

[28] K. O. Elish and M. O. Elish, "Predicting defect-prone software modules using support vector machines," *J. Syst. Softw.*, vol. 81, no. 5, pp. 649–660, May 2008.

[29] Y. W. Kang, J. Li, G. Y. Cao, H. Y. Tu, H. Li, and H. Yang, "Dynamic temperature modeling of an SOFC using least squares support vector machines," *J. Power Sources*, vol. 179, no. 2, pp. 683–692, May 2008.

[30] L. Cao and F. E. H. Tay, "Financial forecasting using support vector machines," *Neural Comput. Appl.*, vol. 10, no. 2, pp. 184–192, May 2001.

[31] W. Hong, "A hybrid support vector machine regression for exchange rate prediction," *Int. J. Inf. Manage. Sci.*, vol. 17, no. 2, pp. 19–32, Jun. 2006.

[32] O. Ivanciuc, "Applications of support vector machines in chemistry," *Rev. Comput. Chem.*, vol. 23, pp. 291–400, 2007.

[33] P. J. H. Hu, T. H. Cheng, C. P. Wei, C. H. Yu, A. L. F. Chan, and H. Y. Wang, "Managing clinical use of high-alert drugs: A supervised learning approach to pharmacokinetic data analysis," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 37, no. 4, pp. 481–492, Jul. 2007.

[34] J. Luo, M. Namburu, K. R. Pattipati, Q. Liu, and S. Chigusa, "Integrated model-based and data-driven diagnosis of automotive antilock braking systems," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 40, no. 2, pp. 321–336, Mar. 2010.

[35] C. D. Katsis, N. Katertsidis, G. Ganiatsas, and D. I. Fotiadis, "Toward emotion recognition in car-racing drivers: A biosignal processing approach," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 38, no. 3, pp. 502–512, May 2008.
[36] V. Vapnik, *The Nature of Statistical Learning Theory*. New York: Springer-Verlag, 1995.
[37] V. Vapnik, *Statistical Learning Theory*. New York: Wiley, 1998.
[38] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining Knowl. Discovery*, vol. 2, no. 2, pp. 121–167, Jun. 1998.
[39] O. Bunke and B. Droge, "Bootstrap and cross-validation estimates of the prediction error for linear-regression models," *Ann. Statist.*, vol. 12, no. 4, pp. 1400–1424, Dec. 1984.
[40] R. R. Picard and R. D. Cook, "Cross-validation of regression models," *J. Amer. Stat. Assoc.*, vol. 79, no. 387, pp. 575–583, Sep. 1984.
[41] D. Ouyang, D. Li, and Q. Li, "Cross-validation and non-parametric k nearest-neighbour estimation," *Econometrics J.*, vol. 9, no. 3, pp. 448–471, Nov. 2006.
[42] D. J. Spitzner, "Constructive cross-validation in linear prediction," *Commun. Statist.—Theory Methods*, vol. 36, no. 5, pp. 939–953, Apr. 2007.
[43] H. Drucker, C. J. C. Burges, L. Kaufman, A. Smola, and V. Vapnik, "Support vector regression machines," in *Advances in Neural Information Processing Systems*. Cambridge, MA: MIT Press, 1996, pp. 155–161.
[44] A. J. Smola and B. Scholkopf, "A tutorial on support vector regression," *Statist. Comput.*, vol. 14, no. 3, pp. 199–222, Aug. 2004.
[45] C. C. Chang and C. J. Lin, LIBSVM—A Library for Support Vector Machines. [Online]. Available: http://www.csie.ntu.edu.tw/~cjlin/libsvm/
[46] K. Rumar, *Infrared Night Vision Systems and Driver Needs*. Warrendale, PA: SAE, Mar. 2003, Tech. Paper: SAE 2003-01-0293.
[47] O. Tsimhoni, J. Bärgman, and M. Flannagan, "Pedestrian detection with near and far infrared night vision enhancement," *Leukos*, vol. 4, no. 2, pp. 113–128, Oct. 2007.
[48] J. H. Lim, O. Tsimhoni, and Y. Liu, "Investigation of driver performance with night vision and pedestrian detection systems: Part 1. Empirical study on visual clutter and glance behavior," *IEEE Trans. Intell. Transp. Syst.*, vol. 11, no. 3, pp. 670–677, Sep. 2010.
[49] D. E. Schmieder and M. R. Weathersby, "Detection performance in clutter with variable resolution," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-19, no. 4, pp. 622–630, Jul. 1983.
[50] H. Chang and J. Zhang, "Evaluation of human detection performance using target structure similarity clutter metrics," *Opt. Eng.*, vol. 45, no. 9, pp. 096 404-1–096 404-7, Sep. 2006.
[51] G. Tidhar, G. Reiter, Z. Avital, Y. Hadar, S. R. Rotman, V. George, and M. L. Kowalczyk, "Modeling human search and target acquisition performance: IV. Detection probability in the cluttered environment," *Opt. Eng.*, vol. 33, no. 3, pp. 801–808, Mar. 1994.
[52] D. L. Wilson, "Image-based contrast-to-clutter modeling of detection," *Opt. Eng.*, vol. 40, no. 9, pp. 1852–1857, Sep. 2001.
[53] T. J. Meitzler, R. E. Karlson, G. R. Gerhart, and E. Sohn, "Wavelet transforms of cluttered images and their application to computing the probability of detection," *Opt. Eng.*, vol. 35, no. 10, pp. 3019–3025, Oct. 1996.
[54] L. Bi, O. Tsimhoni, and Y. Liu, "Using image-based metrics to model pedestrian detection performance with night-vision systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 10, no. 1, pp. 155–164, Mar. 2009.

Luzheng Bi (M'08) received the Ph.D. degree in mechanical engineering from the Beijing Institute of Technology, Beijing, China, in 2004.

He was a Visiting Scholar with the Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor. He is currently an Associate Professor with the School of Mechanical Engineering, Beijing Institute of Technology. His research interests include intelligent human–machine systems, human performance and cognitive modeling, sensing and detection techniques, and driving safety.

Dr. Bi has been a Reviewer for the IEEE Transactions on Intelligent Transportation Systems and the IEEE Transactions on Systems, Man, and Cybernetics. He is an author of refereed journal articles in the IEEE Transactions on Intelligent Transportation Systems, the IEEE Transactions on Systems, Man, and Cybernetics, the *International Journal of Human–Computer Interaction*, and other journals. He received the outstanding Ph.D. dissertation from the Beijing Institute of Technology in 2004.

Omer Tsimhoni (S'99–M'04) received the M.S.E. and Ph.D. degrees in industrial and operations engineering from the University of Michigan, Ann Arbor, in 1997 and 2004, respectively.

From 2005 to 2008, he was an Assistant Research Scientist with the Human Factors Group, Transportation Research Institute, University of Michigan, where he was also an Adjunct Assistant Professor with the Department of Industrial and Operations Engineering. He is currently the Human–Machine Interface Group Manager with the General Motors Advanced Technical Center—Israel, Herzliya, Israel. His research interests include transportation human factors, speech interfaces, adaptive interfaces, computational cognitive modeling of driving, driving safety, and automotive night vision systems.

Dr. Tsimhoni is a member of the Association for Computing Machinery, the Human Factors and Ergonomics Society, and the Society of Automotive Engineers. He is an author of refereed journal articles in the IEEE Transactions on Intelligent Transportation Systems, *Human Factors*, the *ACM Transactions on Computer–Human Interaction*, the *International Journal of Speech Technology*, and other journals. Since 2001, he has consistently presented his work and served as a Reviewer at several conferences, such as the Human Factors and Ergonomics Annual Meeting, Driving Assessment, and the Society of Automotive Engineers Annual Meeting. He was the recipient of several awards for outstanding oral presentations and outstanding student papers.

Yili Liu (S'90–M'91) received the M.S. degree in computer science and the Ph.D. degree in engineering psychology from the University of Illinois at Urbana–Champaign, Urbana.

He is currently an Arthur F. Thurnau Professor and a Professor of industrial and operations engineering with the Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor. He is a coauthor of a human factors textbook entitled *An Introduction to Human Factors Engineering* (Prentice Hall, 1997 and 2003). His research interests include cognitive ergonomics, human factors, computational cognitive modeling, and engineering esthetics.

Dr. Liu is a member of the Association for Computing Machinery, the Human Factors and Ergonomics Society, the American Psychological Association, and Sigma Xi. He is the author of numerous refereed journal articles in the IEEE Transactions on Intelligent Transportation Systems, the IEEE Transactions on Systems, Man, and Cybernetics, the *ACM Transactions on Computer–Human Interaction*, *Human Factors*, *Psychological Review*, *Ergonomics*, and several other journals. He was a recipient of the University of Michigan Arthur F. Thurnau Professorship Award (selected by the Provost and approved by the Regents of the University of Michigan), the College of Engineering Education Excellence Award, the College of Engineering Society of Women Engineers and Society of Minority Engineers Teaching Excellence Award (twice), and the Alpha Pi Mu Professor of the Year Award (five times).