Multimodal Perception and Adaptive Control in Intelligent Robotic Fabrication Systems

Gianfranco Cundari
Department of Computer Science, George Mason University, Fairfax, VA, USA.
gianfranco1978@gmu.edu

Mona Weinhage
Department of Computer Science and Engineering, University at Buffalo, Buffalo, NY, USA.
mona.weinhage183@buffalo.edu

Abstract

Robotic fabrication systems increasingly operate in unstructured or semi-structured environments where material behavior, tool dynamics, and environmental conditions vary during execution. Conventional preprogrammed toolpaths and single-modality sensing are poorly suited to this setting because they cannot resolve the coupled uncertainty that arises when a robot manipulates compliant, partially cured, or geometrically variable materials. This manuscript develops a methodological framework for multimodal perception and adaptive control in intelligent robotic fabrication. The central contribution is an uncertainty-aware architecture that fuses vision, force, pose, and process observations into a fabrication state estimate, propagates that estimate through a Gaussian-process material model, and updates a constrained adaptive control law in real time. The framework treats fabrication as a partially observable dynamic process rather than a static path-following problem. It specifies the fabrication state, the multimodal observation model, a recursive Bayesian estimator, a computed-torque adaptive controller with parameter adaptation, a covariance propagation relation, and a risk-sensitive decision rule for selecting control actions under material and sensing uncertainty. Analytical discussion clarifies stability conditions, calibration requirements, failure modes, and trade-offs between model fidelity and responsiveness. The framework is intended to support robotic systems that must maintain geometric tolerance and process quality when material response cannot be fully characterized before execution. Its significance lies in connecting perception, material inference, and control through explicit uncertainty terms, thereby providing a technically grounded basis for intelligent fabrication systems that adapt during operation rather than after error accumulation.

Keywords

multimodal perception; adaptive control; robotic fabrication; uncertainty quantification; sensor fusion; cyber-physical construction

1. Introduction

Robotic fabrication has expanded from repetitive pick-and-place tasks toward architectural, additive, subtractive, and hybrid processes in which the robot interacts with materials whose state changes during execution [1]. In architectural robotics, manipulators are used for incremental assembly, extrusion, milling, and composite layup, often on scales where tool deflection, material slump, and fixturing error are significant [2]. The dominant engineering assumption in many such systems is that a validated digital model can be converted into a toolpath whose execution is sufficiently repeatable. This assumption is

fragile when the material is heterogeneous, when the tool engages a workpiece with unknown compliance, or when the robot must compensate for drift introduced by vibration, thermal effects, or incomplete calibration. The resulting problem is not simply one of better trajectory tracking; it is one of estimating a changing fabrication state and adjusting control action while the process is still active.

Multimodal perception offers a route to this problem because no single sensor directly observes the quantities that matter. Machine vision can measure geometry, deposition width, surface defects, and tool position relative to fiducials [3]. Force and torque sensing can reveal contact compliance, insertion resistance, and unexpected collisions. Joint encoders and external tracking provide kinematic state, while process sensors such as temperature, pressure, and flow rate indicate material condition. Multisensor fusion improves robustness when individual modalities are noisy or intermittently unavailable [4], but fusion alone does not specify how material behavior should be represented or how control should respond to residual uncertainty. Gaussian-process models provide a useful language for learning and representing spatially varying material response with calibrated uncertainty [5]. Computational material approaches further suggest that material behavior can be treated as an information-bearing variable rather than a fixed input to fabrication planning [6]. Yet these ingredients have not been fully integrated into a control architecture that respects both fabrication tolerances and real-time execution constraints.

A second gap concerns adaptation. Data-efficient learning can reduce the number of trials needed to model a robotic process [7], but learning is not equivalent to control. A fabrication system must decide when to trust a learned material model, when to rely on direct force feedback, and when to slow or stop because uncertainty exceeds a safe or tolerance-compatible bound. The central contribution of this manuscript is therefore a methodological framework that couples multimodal state estimation, Gaussian-process material inference, and constrained adaptive control for robotic fabrication. The framework is analytical rather than experimental: it defines the state and observation models, derives control and adaptation relations, and examines stability, calibration, and trade-off conditions that follow from the formulation. The intention is to provide a coherent design logic for systems that must maintain geometric and process quality under uncertain material response, without treating perception, learning, and control as separate modules.

2. Problem Formulation and System Architecture

A robotic fabrication system can be represented as a controlled dynamical process in which the robot, tool, workpiece, and material state co-evolve. The robot contributes joint positions and velocities, the tool contributes contact forces and process inputs, and the material contributes geometric and mechanical state variables that are only partially observable. This formulation follows the probabilistic robotics view that state estimation and action selection must be treated jointly under uncertainty [9]. It also aligns with simultaneous localization and mapping principles when the workpiece or environment is not known exactly and must be estimated relative to the tool [8]. The fabrication state at time t is defined as a tuple containing robot configuration, tool motion, process variables, material state, and accumulated error.

$$x_t = (q_t, \dot{q}_t, \xi_t, m_t, e_t)$$

Here q_t denotes joint positions, q̇_t joint velocities, ξ_t the tool pose and twist in the workpiece frame, m_t a vector of material or process variables such as temperature,

deposition rate, layer height, or cure state, and e_t the accumulated geometric deviation from the reference fabrication plan. The state is not assumed to be directly measurable. Instead, observations z_t are generated by vision, force-torque, kinematic, and process sensors. The observation model is deliberately left general because different fabrication processes expose different sensing channels, but the estimator requires that each channel provide either a direct measurement or a likelihood function relating observed values to the state.

The fusion problem is expressed through recursive Bayesian estimation. Given observations up to time t, the posterior over the fabrication state is updated from the prior predicted by the process model. This relation is standard, but its terms carry fabrication-specific meaning: the transition model includes robot dynamics and material evolution, while the observation model includes sensor geometry, calibration parameters, and occlusion or contact conditions.

$$p(x_t \mid z_\{1{:}\,t\}) \;\propto\; p(z_t \mid x_t) \int p(x_t \mid x_\{t-1\})\, p(x_\{t-1\} \mid z_\{1{:}\,t-1\})\, dx_\{t-1\}$$

In this expression, p(x_t | z_{1:t}) is the posterior state distribution, p(z_t | x_t) is the observation likelihood, p(x_t | x_{t−1}) is the state transition density, and p(x_{t−1} | z_{1:t−1}) is the prior posterior from the previous step. The integral is over the previous state space. The relation contributes a principled way to combine heterogeneous evidence, but it also exposes a practical constraint: if the transition or observation model is misspecified, the posterior can become overconfident. The architecture therefore separates model estimation from controller execution so that uncertainty can be monitored and used in the control decision.

The system architecture contains four coupled layers. The sensing layer acquires calibrated observations and timestamps them in a common frame. The estimation layer maintains the fabrication state distribution and updates a material-response model when new evidence arrives. The control layer computes tool motion and process inputs using the current estimate and its uncertainty. The supervision layer enforces fabrication constraints, detects inconsistency between predicted and observed behavior, and triggers fallback actions such as reducing speed, changing contact force, or requesting recalibration. This separation is not merely organizational. It allows the adaptive controller to remain stable under bounded estimation error while the material model continues to update at a slower rate.

3. Multimodal Perception and State Estimation

The perception problem in fabrication differs from generic object recognition because the quantities of interest are often continuous, spatially distributed, and process-dependent. Vision can estimate deposited geometry or tool position, but its accuracy falls under specular reflection, occlusion by the tool, or poor contrast between material and substrate. Force-torque sensing can estimate contact state and compliance, but it is affected by tool inertia and unmodeled payloads. Joint encoders are precise but do not capture workpiece deflection. A multimodal estimator must therefore represent both the state and the reliability of each channel. This motivates a likelihood that can be weighted by sensor-specific noise models and by contextual gating, such as discarding vision when the camera is occluded or down-weighting force data during free-space motion.

For material behavior, a Gaussian-process formulation provides a flexible prior over spatially varying response. Let s denote a normalized position or process coordinate within the workpiece, such as a layer coordinate, bead centerline position, or local toolpath parameter. The material response m(s), for example deposition height or springback, is

modeled as a Gaussian process with mean function μ(s) and covariance function k(s,s′). The squared-exponential covariance encodes smooth spatial variation and provides a calibrated variance away from observed locations.

$$m(s) \sim GP(\mu(s), k(s,s')), k(s,s') = \sigma_f^2 \exp(-\|s-s'\|^2/(2\ell^2))$$

Here σ_f² is the signal variance, ℓ is a length-scale controlling spatial smoothness, and ‖s−s′‖ is the distance between two process coordinates. The mean function μ(s) can encode a nominal process model, while the covariance describes residual uncertainty. This model is appropriate because fabrication errors often vary smoothly along a toolpath but are not adequately captured by a single global offset. It also supports data-efficient updating, since the posterior variance can be computed without requiring a dense training set. The model does not replace physical reasoning; it represents the residual after nominal process knowledge has been applied.

The estimator combines the Gaussian-process material model with the recursive Bayesian update. When a new observation arrives, the predicted material state is compared with the measured value, and the posterior over m_t is updated. The controller receives both the posterior mean and the posterior covariance. The mean provides the best current estimate for feedforward compensation, while the covariance determines how aggressively the controller may act. If the variance is large, the controller should reduce speed or force to limit the consequences of a wrong estimate. If the variance is small, the controller can track the planned path more closely. This coupling between estimation uncertainty and control authority is the main mechanism by which perception improves fabrication quality rather than merely documenting error.

Adaptive control theory supplies the formal basis for adjusting controller parameters when the plant or process changes [10]. In the present setting, adaptation has two distinct targets. The first is robot-side uncertainty, such as payload variation, friction, or tool compliance. The second is process-side uncertainty, such as material stiffness or deposition rate. These targets operate on different time scales. Robot-side parameters may be updated continuously, whereas material parameters may require slower updates because each observation is spatially correlated with previous ones. The architecture therefore uses separate adaptation rates rather than forcing all uncertainty into a single estimator.

4. Adaptive Control and Fabrication Planning

The control layer must track a reference tool motion while maintaining contact or deposition conditions and respecting constraints. A common starting point is the computed-torque law for a rigid manipulator, extended with estimated dynamic terms and feedback gains. The robot dynamics and control structure are well established in robotics [11], and nonlinear stability analysis provides conditions under which tracking error remains bounded [12]. In fabrication, the important modification is that the feedforward terms depend on the estimated material state and not only on the planned trajectory.

$$\tau = \hat{M}(q)\ddot{q}_d + \hat{C}(q,\dot{q})\dot{q}_d + \hat{g}(q) + K_p\, e + K_d\, \dot{e}$$

In this equation, τ is the commanded joint torque, M̂(q) is the estimated inertia matrix, Ĉ(q,q̇) contains estimated Coriolis and centrifugal terms, ĝ(q) is the estimated gravity vector, q_d is the desired joint trajectory, and e = q_d − q is the joint tracking error. K_p and K_d are symmetric positive-definite feedback gain matrices. The hats indicate that dynamic terms are estimates and may be updated. The equation contributes a control structure in which feedback

corrects tracking error while feedforward terms shape the nominal motion. In fabrication, q_d itself may be adjusted by the supervision layer when material uncertainty requires a change in speed, layer height, or contact force.

Parameter adaptation can be introduced for uncertain terms that appear linearly in the dynamics or process model. Let $\hat{\theta}$ be the vector of estimated parameters and Y be the regressor matrix evaluated along the measured trajectory. A standard adaptation law updates the parameter estimate in the direction that reduces tracking error.

$$\dot{\hat{\theta}} = -\Gamma Y^T e, \Gamma = \Gamma^T > 0$$

Here $\dot{\hat{\theta}}$ is the parameter update rate, Γ is a positive-definite adaptation gain matrix, Y^T is the transpose of the regressor, and e is the tracking error. The law is intentionally simple, but its use in fabrication requires care. If the regressor is persistently excited only during certain motions, parameter estimates may converge slowly or drift when the robot executes low-acceleration paths. The supervision layer must therefore distinguish between identifiable robot parameters and process parameters that are only weakly observable from the current trajectory.

Planning under uncertainty requires a decision rule that does not reduce to deterministic path following. The controller should choose actions that trade expected fabrication cost against the cost of remaining uncertainty. Let c(x_t,u) be an instantaneous cost representing geometric error, process deviation, and actuator effort, and let λ be a positive scalar that penalizes the trace of the predicted state covariance. The trace is a scalar summary of total uncertainty, and its inclusion encourages actions that reduce or avoid uncertainty when the expected cost alone would not.

$$u_t* = argmin_u\, E[c(x_t, u) + \lambda\, tr(\Sigma_\{t+1\}) \mid z_\{1{:}t\}]$$

In this expression, u_t is the selected control input, E is the expectation under the current state posterior, Σ_{t+1} is the predicted covariance at the next step, and tr(·) is the matrix trace. The rule is not a substitute for safety constraints; it is a ranking criterion among feasible actions. It explains why an intelligent fabrication system may deliberately slow down or take a sensing pass even when the nominal trajectory is feasible: the reduction in uncertainty can lower the expected cost of later decisions.

5. Uncertainty, Stability, and Trade-offs

Uncertainty propagation is central because perception errors do not remain confined to the estimator. They enter the control law through estimated parameters, through the reference trajectory, and through constraint evaluation. For an extended Kalman filter approximation, the state covariance can be propagated using a linearized transition model. The following relation captures the main terms that must be monitored.

$$\Sigma_t = F_t \Sigma_{t-1} {F_t}^T + Q_t - K_t S_t {K_t}^T$$

Here Σ_t is the posterior covariance, F_t is the linearized state-transition matrix, Q_t is the process-noise covariance, K_t is the estimator gain, and S_t is the innovation covariance. The equation is not merely a bookkeeping device. It identifies the three ways uncertainty changes: process noise injects new uncertainty, linearized dynamics transport existing uncertainty, and measurement update reduces uncertainty through the innovation term. If Q_t is set too low, the estimator may reject legitimate material variation as measurement noise. If Q_t is set too high, the controller becomes conservative without gaining accuracy. Calibration of Q_t therefore has direct fabrication consequences.

Stability analysis must account for the fact that the controller uses estimated states and parameters. If the estimation error is bounded and the adaptation gain is sufficiently small relative to the control bandwidth, the tracking error can remain uniformly ultimately bounded under standard assumptions on the dynamic terms. This conclusion depends on smoothness, bounded disturbances, and the absence of unmodeled contact discontinuities. In fabrication, contact can change abruptly when the tool enters or leaves the material, so the controller must include a supervision layer that detects contact transitions and switches between free-motion and contact-mode gains. The adaptive law should not be allowed to update aggressively during these transitions, because the regressor may change discontinuously and drive parameter estimates away from their physical range.

Fabrication constraints are often expressed as inequalities on tool position, force, temperature, or material flow. For a set of constraint functions g_j(x_t,u_t), a feasibility margin can be defined using the signed distance to the nearest active constraint. This margin provides a scalar measure of how close the current state is to violating a process limit.

$$\gamma_t \;=\; min_j \; g_j(x_t, u_t) \;/\; \|\nabla g_j(x_t, u_t)\|$$

In this relation, γ_t is the minimum normalized margin, g_j is the j-th constraint function, and ∇g_j is its gradient with respect to the state and input. A positive γ_t indicates that the state is inside the feasible region, while a small positive value indicates proximity to a boundary. The normalization by the gradient makes margins comparable across constraints with different units. The controller can use γ_t to reduce speed or force when the margin is small, even if the nominal tracking error is acceptable. This is important in fabrication because a small geometric error may be tolerable, but a small margin to collision or material collapse may not be.

There is a fundamental trade-off between model fidelity and responsiveness. A richer material model can reduce prediction error, but it also increases estimation and computation time. A highly adaptive controller can compensate for unmodeled effects, but it may respond to sensor noise or transient material irregularities. The framework addresses this trade-off by assigning different time scales to material inference and motion control. Material inference updates at a rate consistent with spatial correlation and process observation, while motion control runs at the robot servo rate. The separation prevents slow material learning from destabilizing fast feedback loops.

6. Technical Discussion and Implications

The proposed framework differs from deterministic robotic fabrication in its treatment of uncertainty as a control variable. In deterministic systems, perception is used to detect errors after they occur, and adaptation is often limited to calibration offsets. In the present formulation, uncertainty affects both the estimated state and the choice of action. This has several implications. First, sensor calibration becomes part of control design rather than a preprocessing step. A camera extrinsic error or force-sensor bias changes the observation likelihood and therefore changes the posterior covariance. Second, the material model must be falsifiable. If observations repeatedly fall outside the predicted covariance, the model should be revised or its process noise increased. Third, the controller must have a defined fallback behavior when uncertainty exceeds a threshold. Slowing down, reducing force, or pausing for additional sensing are legitimate control actions, not failures of autonomy.

The framework also clarifies why black-box learning alone is insufficient for fabrication. A learned policy can map observations to actions, but it does not by itself expose the

uncertainty that determines whether an action is safe. Gaussian-process or Bayesian representations provide explicit uncertainty, but they require assumptions about smoothness and correlation. The practical choice is therefore hybrid: use physically motivated nominal models for the dominant dynamics, use probabilistic models for residual spatial variation, and use feedback control to maintain stability when both are imperfect. Additive manufacturing process models illustrate this need because layer-wise geometry and thermal history introduce path-dependent uncertainty that cannot be captured by a single global correction [17]. Compliant and soft material systems further show that material deformation can be part of the fabrication mechanism rather than an error to be eliminated [16]. The controller must therefore distinguish between acceptable compliance and unacceptable deviation.

The architecture also has boundary conditions. It assumes that sensors are time-synchronized and calibrated to a common frame. It assumes that the robot can execute torque or position commands with sufficient bandwidth to realize the computed control action. It assumes that material behavior varies smoothly enough for the chosen covariance function over the relevant spatial scale; if the process exhibits abrupt phase changes or discrete failures, a Gaussian-process model may be inappropriate without a change-point or switching extension. The framework does not remove the need for process-specific experimentation. It instead defines what must be measured and how those measurements enter the control decision.

From an implementation perspective, the most demanding element is not the control law but the consistency of the state representation. Vision measurements may be expressed in camera coordinates, force measurements in tool coordinates, and material measurements in workpiece coordinates. The estimator must transform all observations into a common frame and propagate the associated Jacobians. If this transformation is neglected, the posterior covariance will be inconsistent even when each sensor is individually accurate. Similarly, the controller must distinguish between uncertainty in the robot state and uncertainty in the material state. The former can often be reduced by better encoders or calibration, while the latter may be irreducible within the current process and must be handled by conservative action selection. Robot modeling and control texts provide the kinematic and dynamic tools for the first problem [15], while probabilistic estimation provides the language for the second [14]. The contribution here is to place both within a fabrication-specific decision loop.

7. Conclusion

This manuscript developed a methodological framework for multimodal perception and adaptive control in intelligent robotic fabrication systems. The central problem is that fabrication state is only partially observable and evolves during execution, so preprogrammed toolpaths cannot maintain quality when material response, contact conditions, or robot dynamics deviate from nominal assumptions. The framework defines a fabrication state that includes robot configuration, tool motion, process variables, material state, and accumulated error. It combines recursive Bayesian estimation with a Gaussian-process material model, computes adaptive control actions using estimated dynamics, and incorporates uncertainty and feasibility margins into action selection. The derived relations show how observation likelihoods, covariance propagation, parameter adaptation, and constraint margins connect perception to control. The analysis also identifies conditions that limit the approach: sensor calibration consistency, smoothness assumptions in material modeling, bounded estimation error for stability, and the need for fallback behavior when uncertainty exceeds safe limits. The practical significance is that fabrication systems can be designed to adapt during

operation while preserving explicit links between what is sensed, what is inferred, and what is commanded. This provides a technically grounded alternative to treating perception, material modeling, and control as separate engineering problems.

References

1. Keating, S. J., Spielberg, N. A., Klein, J., & Oxman, N. (2017). Toward site-specific and self-sufficient robotic fabrication on architectural scales. Science Robotics, 2(5), eaam8986.
2. Brell-Çokcan, S., & Braumann, J. (Eds.). (2013). Robotic fabrication in architecture, art and design 2012. Springer.
3. Perez, L., Rodríguez, Í., Rodríguez, N., Usamentiaga, R., & García, D. F. (2016). Robot guidance using machine vision techniques in industrial environments: A comparative review. Sensors, 16(3), 335.
4. Khaleghi, B., Khamis, A., Karray, F. O., & Razavi, S. N. (2013). Multisensor data fusion: A review of the state-of-the-art. Information Fusion, 14(1), 28-44.
5. Rasmussen, C. E., & Williams, C. K. I. (2006). Gaussian processes for machine learning. MIT Press.
6. Luo, D., Chen, D., Wang, J., Zhu, G., & Xu, W. (2020). The smart robot crafting approach to computing materials. Construction Robotics, 4(3), 239-249.
7. Deisenroth, M. P., Fox, D., & Rasmussen, C. E. (2015). Gaussian processes for data-efficient learning in robotics and control. IEEE Transactions on Pattern Analysis and Machine Intelligence, 37(2), 408-423.
8. Cadena, C., Carlone, L., Carrillo, H., Latif, Y., Scaramuzza, D., Neira, J., Reid, I., & Leonard, J. J. (2016). Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age. IEEE Transactions on Robotics, 32(6), 1309-1332.
9. Thrun, S., Burgard, W., & Fox, D. (2005). Probabilistic robotics. MIT Press.
10. Slotine, J.-J. E., & Li, W. (1991). Applied nonlinear control. Prentice Hall.
11. Siciliano, B., & Khatib, O. (Eds.). (2016). Springer handbook of robotics (2nd ed.). Springer.
12. Khalil, H. K. (2002). Nonlinear systems (3rd ed.). Prentice Hall.
13. Hall, D. L., & Llinas, J. (1997). An introduction to multisensor data fusion. Proceedings of the IEEE, 85(1), 6-23.
14. Ko, J., & Fox, D. (2009). GP-BayesFilters: Bayesian filtering using Gaussian process prediction and observation models. Autonomous Robots, 27(1), 75-90.
15. Siciliano, B., Sciavicco, L., Villani, L., & Oriolo, G. (2010). Robotics: Modelling, planning and control. Springer.
16. Rus, D., & Tolley, M. T. (2015). Design, fabrication and control of soft robots. Nature, 521(7553), 467-475.
17. Gibson, I., Rosen, D. W., & Stucker, B. (2015). Additive manufacturing technologies: 3D printing, rapid prototyping, and direct digital manufacturing (2nd ed.). Springer.